

Health reform: reality strikes USA . . .

US President Bill Clinton's proposal for health-care reform is in for tough times now that it is being debated in Congress, because no one is entirely happy with the legislative details.

FROM the start, no one (except perhaps the president and his wife) believed that Congress would turn Bill Clinton's health-care reform plan into legislation intact. But now that the Clinton plan, in all its complexity, is actually before the Congress for debate, it looks as if it may be torn to shreds. This is not necessarily a bad thing.

Although observers were aware of the nature of the plan from the start, now that Congress is taking it apart line-by-line it is even clearer that the Clinton plan would create a monstrous new bureaucracy. Hilary Rodham Clinton has said in defence that the plan is no more bureaucratic than the present health-care system, but that is not a good argument in favour of either one.

The chief virtue of the Clinton plan, largely crafted by Mrs Clinton and Ira Magaziner (a management guru whose experience in health-care is close to zero), is that it provides "universal coverage". That one of the wealthiest nations in the world, with the most sophisticated of modern medicine, should offer coverage to everybody is a goal long overdue. Indeed, universal coverage is integral to most of the competing pieces of draft legislation before Congress. But the real issues are financing and logistics.

At present, the really poor are covered by the government's Medicaid programme. And the vast majority of people, from lower to upper class, have adequate insurance through their employers. It is the so-called "working-poor" (people with part-time or low-income jobs with small companies, for instance) who comprise the majority of the 37 million uninsured. In order for these people to receive government-mandated coverage, someone will have to subsidize the cost of their insurance premiums. Clinton's plan puts the burden for at least 80 per cent of the cost on employers who would be required to offer insurance.

But small businesses are claiming loudly (and not without merit) that mandated insurance payments would force them out of business, thereby simply throwing the working-poor out of work. As for big businesses, they have been subsidizing medical care for years. In a perfectly legal but publicly little-understood system, hospitals place a kind of surcharge on insured patients in order to cover the costs of the poor. Medicaid, for instance, does not reimburse hospitals for full costs. In short, the insured have been supporting the poor and uninsured all along.

But would the Clinton plan not change that? In principle, yes. By guaranteeing universal coverage, it requires that everyone would have insurance. But since not everyone can

pay, or pay equally, for that insurance, some form of subsidy will still be required. A good case can be made for facing the truth that the subsidy will have to come from taxes. But neither the administration nor Congress is quite there yet.

The other feature of the Clinton plan that spells trouble is Magaziner's devotion to the notion that all of America (perhaps like Gaul) can be divided into distinct parts known as regional health alliances. These would be brand-new bureaucracies that would collect premiums and make sure that each alliance offered a choice of health-care plans — good, better and best.

This scheme has run into vehement opposition. First, Americans are a mobile lot. Life under a regional alliance provides daunting bureaucratic obstacles to the person from New York who gets sick in San Francisco. The cost-cutting, deal-making nature of the alliances would virtually eliminate the current freedom to travel for care of particularly difficult medical problems. It is not just the wealthy who go to Houston or Cleveland for cardiac surgery. It is not just the well-connected who fly to Pittsburgh for organ transplantation. It would take a powerful argument indeed to convince a penny-pinching alliance manager to permit patients to go out of their region to get what they believe to be the best, rather than merely adequate, care. So, the majority of people who have good insurance coverage now will lose under the Clinton plan with its greater bureaucracy and tendency to equate quality with the lowest common denominator.

Is health care reform needed? Yes. Is universal coverage an appropriate goal? Yes, both morally and politically. Has Clinton got it right? Not yet. It will be a long spring and summer as Congress and the administration struggle to produce a plan that is just, consistent with freedom of choice and movement, and also economically feasible. □

. . . as UK faces the bill

The British government needs to redress its political neglect of health-related research.

MEDICAL research, despite its prominence among Britain's scientific achievements, has taken a back seat in recent manoeuvrings. Sometimes this has been beneficial; the government's primary concern with wealth creation has left the Medical Research Council (MRC) virtually intact. But in the long run it cannot be good that research has been given



scant attention in the recent restructuring of the National Health Service.

That neglect highlights an uncomfortable mismatch between the government's approach to the "rationalization" of state support and the optimal distribution of public resources in support of advanced research. Nowhere are the dangers of this gap more vivid than in the current controversy over the possible closure of the Hammersmith Hospital in west London, and the transfer of both patient care and research — including in particular the activities of the Royal Postgraduate Medical School and the MRC's new Clinical Science Centre — to the nearby Charing Cross Hospital in Fulham.

The economic case for a merger, made by the Tomlinson report on the provision of health-care in London, is strong. With changes both in the population of their respective catchment areas, and in the procedures by which patients are referred to hospitals by physicians outside London, neither hospital is likely to remain viable.

It is also true that the research strengths of the two institutions are sufficiently complementary for a merger to make sense. Indeed, the elements now exist in west London to build up a multi-campus medical institution with an international status.

But two factors threaten such an imaginative step. The first is parsimony on the part of the Treasury. As the report delivered to the Health Department last week points out (see page 176), moving the research facilities at the Hammersmith to the Charing Cross site will be expensive. In order to demonstrate its genuine commitment to the move, the government needs both to confirm that this money will be made available also to ensure that it cannot be subsequently withheld by a new Chancellor of the Exchequer.

The second factor is political. Exchanges in the House of Commons have already suggested that Mrs Virginia Bottomley may be more sympathetic to demands from the local Conservative Member of Parliament to keep Charing Cross open (and move the Hammersmith Hospital on to its site) than to those for a move in the opposite direction. Fulham is a marginal Tory constituency, and Hammersmith a safe Labour stronghold; but it would be scandalous if party politics were to dominate a decision likely to affect the health and well-being of so many.

What should the government do? Doing nothing is still an option — both hospitals could remain open and operate independently — but it would be the worst of all worlds. Uncertainty would remain, and many of the researchers at both institutions would start looking for more secure situations, a recipe for stagnation and eventual collapse.

The arguments about which direction a merger should move in remains finely balanced. The price for the institution required to vacate its existing premises will obviously be high. The important thing at this stage is that the government rejects the temptation of short-term gains — economic or political — and evolves a strategy for developing the unique bond of high-quality health-care with both clinical and basic research for which Hammersmith has built up an enviable reputation. Anything short of a full-blooded commitment to this goal will be a recipe for disaster. □

British beef again

The possible German ban on British beef cannot be brushed aside, but needs European action.

THE proposal by the German health authorities last week to ban imports of British beef will be a serious blow to the British beef industry. For the fuss is yet another reminder that British cattle are infected with bovine spongiform encephalopathy (BSE), the cow's equivalent of scrapie in sheep and Creutzfeldt-Jakob disease (CJD) in people. The infection has been established for a decade. Since 1988, there have been strict controls on the feeding of sheep offal to cattle, the presumptive cause of the original infection of British cattle herds, but the infection has not gone away. Now, the German authorities have declared that Germans no longer wish to risk contracting CJD from contaminated beef. The British authorities correctly say that there is no evidence that BSE leads to CJD. But equally, there is no evidence that BSE cannot cause CJD. There cannot be.

The stand-off on BSE is thus both a classic illustration of the difficulties of dealing with inponderable risks to public health and a test of what the European Union really means. The technical dilemma is haunting. Only in the past year or so has the mechanism of scrapie, BSE and CJD become evident in outline; the details remain obscure. The infectious particles are aberrant forms of a naturally occurring protein molecule, and should not survive the mammalian stomach, but transmission experiments in the past few years have shown that mink, rodents and ungulates such as deer can be infected by the oral route. The British regulations on cattle feed have been rigorously enforced, animals with overt symptoms have been slaughtered and the Medical Research Council has established (at Edinburgh) a unit to monitor CJD in British people. (There has been a statistically insignificant increase in recorded cases to about 40 a year, probably a sign of more vigilant diagnosis.) The worry now is that cases of BSE are cropping up in animals born since 1988, suggesting that maternal transmission to calves is a reality. The figures for 1993 will crucially indicate whether continued regulation of cattle feed will suffice to contain BSE. Nobody should be surprised if it will not.

The risks that eating infected beef will cause CJD are strictly hypothetical. There is no evidence of a case of CJD caused by BSE, but the transmission of BSE to other mammals equally raises the question whether people can be immune. Why should Germans run even a hypothetical risk when there is plenty of uncontaminated beef on the European market?

To restore faith in British beef, British veterinarians will have to convince their opposite numbers on the mainland, offering them open access to such data as there are. They should also begin preparing to replace the British cattle populations with animals free from the infection. The cost would be enormous — comparable with the £30 billion or so annual cost of Europe's Common Agricultural Policy. Either way, Britain will need the rest of Europe's help. □

French geneticists split over terms of commercial use of DNA bank

Paris. A proposed deal that would have given Millennium, a Massachusetts-based biotechnology company, exclusive access to a bank of DNA collected from over 800 French families has triggered a fierce debate in France over how efforts to sequence the human genome project should be commercialized.

The controversy centres on claims by Philippe Froguel, a researcher at the Centre d'Etudes du Polymorphisme Humain (CEPH) in Paris, that he has lost his job at the centre because of his opposition to the proposed deal. His group developed the DNA bank with samples taken from more than 5,000 individuals, and also discovered the first gene for non-insulin-dependent diabetes in 1992 (*Nature*, **356**, 162; 1992).

Froguel says that CEPH had been about to provide Millennium with exclusive access to the bank for \$400,000 a year for four years, a relatively small fraction of his laboratory's FFr 8.2 million (US\$1.8 million) annual budget.

He also argues that Daniel Cohen, the director of CEPH, is facing a conflict of interest, as he is also both a founding scientific director and a shareholder of

Millennium. Froguel says that Millennium needs access to the bank to secure a broad alliance with the Swiss company, Hoffmann La Roche, aimed at developing drugs to treat diabetes.

But Cohen says that he has left all decisions about the proposed Millennium contract to the board of CEPH, which includes government representatives as the centre is partly financed by the government. Mark Levin, chief executive officer of Millennium, says the deal being negotiated with Roche is a broad one, and does not depend on having access to DNA from the French families.

Cohen says that both he and the board of CEPH rejected Millennium's initial proposal on the grounds that it was inequitable, that it did not make clear the terms of CEPH's scientific participation, and that it may have violated French ethical regulations. But he says the board has told him to continue negotiations in the hope of reaching an agreement, and he has proposed that the CEPH bank be made available to users for a fee.

Froguel claims that the validity of his concerns were confirmed last week by the intervention by Alain Pompidou, scientific

adviser to the prime minister Edouard Balladur. Pompidou says that the contract under discussion between the French centre and the US company did not provide for a fair partnership. But he says the government has not blocked a deal.

The government has now set up a committee to decide how the commercial use of DNA data-banks should be regulated. Indeed, the conflict at CEPH has precipitated a broader debate on the use of information about the human genome.

Many French researchers were annoyed at the prospect that Millennium would have obtained exclusive access to the DNA bank. But one leading US genome researcher argues that companies such as Millennium have little interest in getting exclusive access to DNA banks. "It's not like a patent issue," he says. "Millennium would get a head start; it would not give them a blocking position."

He says that diabetes research requires access to DNA banks from different ethnic groups, and that companies realize that, if they start privatizing DNA banks, they will all lose. Levin says the wariness of French researchers about commercial ties is reminiscent of the early days of biotechnology in the United States.

He also points out that, as there is no open access to CEPH's bank, CEPH and the French government are already effectively exercising an exclusive licence on it. Axel Kahn, director of the INSERM laboratory of Genetics and Molecular Pathology in Paris, says that this is a general problem where research could potentially lead to important applications. A solution, he suggests, would be to oblige researchers to make their resources freely available after a reasonable period in which to exploit them themselves.

Jean-Francois Mattei, the rapporteur for three bioethics bills currently before the French Parliament, claims says that the commercial value of DNA banks had not been anticipated in drawing up current guidelines. As a result, subjects only gave their informed consent to the use of their DNA for research, not for profit.

As a result of his public dispute with Cohen, Frogeau has now been asked to leave CEPH by its president, Jean Dausset. At the same time, Bernard Barateau, president of the French Muscular Dystrophy Association, has decided to withdraw the association's FFr3 million annual funding of diabetic research at CEPH. "We are interested in DNA being freely available to researchers worldwide," says Barateau. "We have no time for such conflict." **Declan Butler**

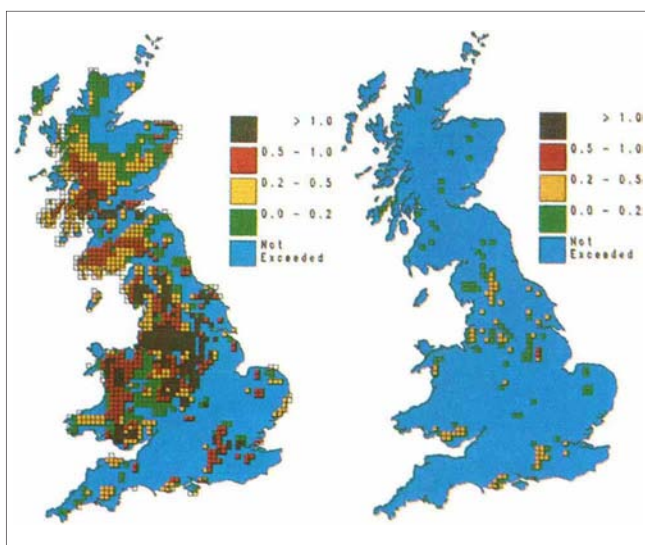
Maps show gains of reduced SO₂ emissions

London. A dramatic illustration of the potential impact of reduced sulphur emissions on acidity levels in the soil has been provided in a report published last week by the Critical Loads Advisory Group of the UK Department of the Environment.

The group used mineralogical information to construct a map showing how much sulphate each region can receive before soil acidification begins, and compared these values to the regional

distribution of atmospheric sulphate deposition in the period 1989–91. The result is a map (left) showing areas that receive more than the critical load of sulphate, and therefore experience acidification.

The group then used this approach to predict the likely effect of attempts to reduce industrial emissions by modelling



future changes in regional sulphate deposition. The second (right) map shows the benefits to be obtained by reducing emissions by 73 per cent from their 1980 levels, in line with the United Nations Economic Commission for Europe's sulphur protocol negotiations.

Gabrielle Walker

UK 'must pay up' for costs of hospital merger

London. The British government has been urged to make an "unequivocal commitment" to provide at least £100 million to secure the future of research at the Hammersmith Hospital in London — one of the country's leading clinical research institutions — if, as many expect, it decides to move the hospital facilities to a site several miles away already occupied by the Charing Cross Hospital.

The proposal is made implicitly in a report prepared for the Department of Health by Sir David Phillips, the former chairman of the Advisory Board for the Research Councils, and Sir Rex Richards, a former vice-chancellor of the University of Oxford, which was presented to Mrs Virginia Bottomley, the health secretary, last week.

The report says that, in order to curtail uncertainty over the fate of the various institutions sharing the Hammersmith Hospital site, a decision on whether the hospital is to be moved should be taken "urgently and irrevocably". These institutions include in particular both the Royal Postgraduate Medical School (RPMS) and the Medical Research Council's (MRC's) Clinical Sciences Centre, due to be opened in April.

But it also concludes that the money needed to ensure the continued quality of Hammersmith's research must both be protected and "seen to be protected" until the completion of any move, which is likely to take a minimum of five years.

"We appreciate that this would be difficult, given the financial framework within which the government operates," say the two authors of the report. "But we believe it to be essential that some means be found to assure researchers that intentions would be implemented, and that funding would not be a cause of delay or change."

The report comes at a time when health department officials are completing final discussions on a decision that will have a major

impact on the future of London as one of the world's major centres of biomedical research.

The merger between the two hospitals has been proposed as a response to falling demand for beds at the two hospitals, itself the result of both demographic shifts and government policies that are discouraging physicians elsewhere in Britain from sending their patients to London for treatment.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Hammersmith Hospital's future is in the balance.

No decision has been taken on which of the two will close. But the Charing Cross Hospital is the more popular of the two with both patients and general practitioners, because it is considerably more accessible than the Hammersmith. It is also situated in a marginal parliamentary constituency currently held by the Conservatives.

Medical researchers at the Hammersmith Hospital, however, are warning that, if their institution is to be the one that has to move, the future of many research teams with an international reputation could be threatened.

Almost a thousand top biomedical researchers have signed a letter to Bottomley pointing out that last year, for example, two Hammersmith research departments achieved top scores in a research assessment exercise carried out by the Higher Education Funding Council for England. They also emphasize that the hospital has been selected by the MRC as one of the main

focal points of a new initiative to boost long-term clinical research in areas such as human genetics and immunology.

Already, however, uncertainty over the future of the Hammersmith is causing difficulties. For example, those responsible for recruiting research teams for the new Clinical Sciences Centre, being funded jointly by the MRC and the Wellcome Trust, are finding some positions difficult to fill (*Nature* 365, 477; 1993).

Many existing staff are considering whether to take up offers of jobs elsewhere. "The problem is that, faced with the prospects of a major upheaval five years down the road, people will just leave," says Karol Sikora, professor of oncology.

Phillips and Richards were not asked to judge whether the merger was a good idea, or even which of the two sites should be the one occupied by the new, single hospital, but to comment on the feasibility of a merger and the steps needed to minimize the impact on research.

The report endorses the long-term goal of grouping institutions such as the RPMS into a multi-campus College of Medicine, bringing together a number of West London biomedical teaching and research institutions under a single administration, with undergraduate teaching focused on Imperial College. It says that, as a first step towards this goal, the government needs to commit itself to funding a new pre-clinical building at Imperial College.

Officials at Hammersmith, still keen to avoid being moved to the Charing Cross Hospital, are emphasizing the endorsement the report provides of the high quality of research at the hospital. "The report points out that to move the Hammersmith to Charing Cross will be expensive and risky," Sir Colin Dollery, the dean of RPMS, said last week. "There is no previous experience of moving a hospital of this complexity, and it would be both unwise and unnecessary."

No firm figures are quoted by Richards and Phillips in their report. But one calculation is that at least £100 million, and probably considerably more, would be needed to reproduce on the Charing Cross site all the research facilities currently in operation at Hammersmith.

So far, the government has given no indication that it will be prepared to guarantee to provide the necessary funds. But as Sir Walter Bodmer, director-general of the Imperial Cancer Research Fund Laboratories, puts it: "If the decision is made for the Charing Cross site without securing the money needed to make the move properly, the Hammersmith will be destroyed, and anyone who makes that decision will rue the day they did so."

David Dickson

Hot-line will answer scientific queries

London. An information line allowing the British public to have their scientific questions answered over the telephone is to be opened in London next week by William Waldegrave, the minister for science, as part of the government-backed Science Week.

Known as Science Line, the information service has already been tested on a limited scale following the broadcast of various television programmes on scientific topics. More than 800 calls, for example, were received last summer following a programme on dinosaurs, and there was a similar response to programmes on memory

and gene therapy.

The expansion of the service is being made possible primarily by a grant from the Wellcome Trust, with additional support from the Royal Society and Channel 4 television, as well as the Office of Science and Technology.

Other activities included in Science Week range from laboratory open days in many universities to an experiment being organized by the British Broadcasting Corporation to test the hypothesis that people lie more convincingly on video than either on radio or in print. □

US universities oppose cap on federal funding

Washington. Universities across the United States are mounting a fierce campaign to kill a provision in President Bill Clinton's budget proposals for next year which, they claim, threatens to throw their finances into chaos whilst unfairly penalizing those institutions that are most successful in attracting research money.

The provision, contained in the administration's 1995 budget proposals published last month, would cap at 1994 levels the total amount of money that any university could claim from the federal government to cover the indirect costs of research.

Finance officers at large research universities are already planning multimillion dollar cuts from this summer to comply with the proposal (see below), while Washington lobbyists and well-connected university presidents are marshalling all the forces they can to kill it in Congress.

The amount of money to be held back is small, between \$130 million and \$170 million, out of the \$12 billion the federal government will pay for university research this year. But universities are terrified of the proposal.

Their main concerns are that what the budget language calls a "one-year pause" in the growth of federal support for indirect costs could become a permanent freeze, and that by reopening the issue of indirect costs, the proposal will invite hostile congressmen to focus on such costs as a way of attacking universities in general.

The universities are also angry that Clinton wrote to the president of Massachusetts Institute of Technology, Charles Vest, last spring, promising that he would not seek to cap indirect costs.

The universities may have lost valuable weeks in fighting the proposal — which they heard of only days before the budget was published — because some feared that a public fight could create a backlash. Indirect costs are soft ground for the universities to fight on, as they cover expenses ranging from lighting, heating and gardening to administrative costs, each of which Congress has been reluctant to pay for.

Last week, however, Cornelius Pings, the president of Association American Universities (AAU), sent an unusually forthright letter to John Gibbons, the president's science adviser, and Alice Rivlin, deputy director of the Office of Management and Budget, describing the proposed move as "bad science policy, bad public policy and flawed budgeting".

Pings said in the letter that the provision would impose "serious damage" on the relationship between government and universities, and warned of an "erosion of credibility, with human and financial consequences which can now only be guessed at, but

which will be neither trivial nor short-term".

He also warned that "yours will be the first administration in fifty years to raise serious questions of dependability and trust" by setting "a rule both arbitrary and gratuitous". It would leave a "clear sense of cavalier underconcern for the impact on the universities and the nation's science programme".

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Pings: warns of serious damage

"We are not whining about the money," he wrote. "If you really need \$130 million, just take it. But please have the forthrightness to take it by adjusting your commitment to total science support." Pings points out that the proposal will inflict the most pain on successful universities, while those whose research programmes have stagnated will be left unscathed.

The AAU says that a freeze on overhead payments would directly undermine universities which have invested money in new research laboratories. Such facilities were often cost-justified on the basis that they

According to Pings, the implementation of the provision would lead to "great stress and game-playing within the program agencies and on the campuses", and universities might choose to sue the government for breach of contract.

"We are not whining about the money," he wrote. "If you really need \$130 million, just take it. But please have the forthrightness to take it by adjusting your commitment to total science support." Pings points out that the proposal will inflict the most pain on successful universities, while those whose research programmes have stagnated will be left unscathed.

The AAU says that a freeze on overhead payments would directly undermine universities which have invested money in new research laboratories. Such facilities were often cost-justified on the basis that they

would attract extra federal research funding, the indirect element of which would pay off the loan on the building. But under the proposed freeze, no additional money could be forthcoming.

The precise way in which the proposal would work has not been finalized. But as it now stands, it would require each university to add together its indirect funding from all federal sources and then ensure that the 1995 figure did not exceed that for 1994. The problem of compiling such figures may get worse when the congressional appropriations subcommittees responsible for different agencies decide how to interpret the proposal.

But some key congressional staff are sympathetic to the AAU's arguments. One who deals with funding for the largest research agency, the National Institutes of Health, says that, although the Clinton proposal in its present form has no support, his committee would be reluctant to replace it with a straight cut in research funding.

The "pause" proposal was accepted as part of the outline budget proposal passed by the House of Representatives last week. But the universities are optimistic that it will be modified as the detailed budget is hammered out in appropriations subcommittees over the summer. Their difficulty is that they are playing a zero-sum game: whatever progress they make in defending their indirect cost base will probably be at the expense of the rest of the science budget.

Colin Macilwain

Los Angeles plans for the worst

The University of Southern California (USC) has made exceptional progress in attracting federal research money in recent years. But the administrative vice-president, Dennis Dougherty, must be starting to ask if it was worth the effort.

USC is a private institution in Los Angeles, outside the public University of California system, and is one of the top 20 US research universities. In the current financial year, which ends on 30 June, it will attract around \$215 million in federal funds, of which \$52.5 million is to cover indirect costs.

Given recent strong growth at the schools of medicine, engineering and basic sciences, Dougherty had expected the total to increase to \$235 million next year, with \$58 million for indirect costs. But he is now busy preparing for a threatened freeze on the indirect portion which, if passed by Congress, will leave the university \$5.5 million short.

Dougherty says the result will be "a

general wringing out" of costs, similar to a previous round of cuts that cost 600 support jobs at USC in 1992. This time, he says, the cuts will "change the quality of life" on his campus in south central Los Angeles.

About half of the savings will be made centrally, he says, leaving the deans at the three main research schools to find the rest. "The deans are going to have less opportunity to recruit, and the impact will roll through into salary policy," he says.

"If they want to save \$150 million they should just appropriate \$150 million less," asserts Dougherty. "What they've done is increase it by 10 per cent, and say they're not going to pay you for light, fuel, buildings or support. The longer-term implication is that if we are going to expand aggressively, we need to be more confident that our partners in government agencies can be depended upon to pay their share." □

Germany uses BSE fears to seek ban on British beef

Munich. German scientists are split over efforts being made by the country's health minister, Horst Seehofer, to secure a Europe-wide ban on the import of British beef on the basis that it could be infected with bovine spongiform encephalopathy (BSE).

The German minister's proposal is based on concern that the disease, which first broke out among cattle in the mid-1980s and causes fatal neurological symptoms similar to those associated with the human Creutzfeldt-Jakob disease (CJD), could be transferred to humans through the consumption of meat.

Last week, acting on a recommendation from the Bundesgesundheitsamt (BGA), the German federal health agency, Seehofer suggested to his fellow European health ministers that the European Commission (EC) impose a ban on the import of British beef by all countries of the European Union (EU).

BGA scientists say the possibility of cross-species disease transfer has not been positively disproved, and that evidence of a few cases of BSE-like syndrome in zoo animals has raised new concerns.

But Hans-Dieter Klenk, a virologist from the Marburg Institute for Virology and one of five independent scientists who took part in a meeting organized by the BGA last December to discuss what action Germany should take in the light of public fears about British beef — and who concluded that there was "no scientific finding that BSE is transferable to humans" — says that a ban is "too early".

Seehofer's actions are based on the fact that the current BSE epidemic in Britain has almost certainly been caused by the practice of feeding cattle with a mix containing the brains of sheep suffering from a similar disease, scrapie (a practice stopped in 1988).

The epidemic is being checked in Britain by slaughtering infected herds, and imposing stringent controls to ensure that all beef intended for consumption is free from nervous tissue and other potentially infected offal. British government scientists claim these controls are sufficient to protect the public. "The facts are clear," says Ray Bradley from Britain's Central Veterinary Laboratory in Weybridge. "There is no evidence anywhere in the world that diseases of the spongiform encephalopathy group can be transferred naturally through meat consumption."

Despite this, Seehofer is now trying to convince his EU partners that they should adopt the German recommendation for a ban on British beef. He and other members of the German cabinet — in a move supported by the German parliament — is now planning to push for agreement in two fur-

ther meetings later this month in Brussels.

Seehofer has also said that if a Europe-wide ban is not imposed by the commission, Germany will impose its own national ban on British beef. But there is a wide perception in Germany that his position has more to do with politics than with science.

Last year, Seehofer ordered the disbanding of the BGA because, he said, it was not keeping him sufficiently informed of important health threats, and in particular had not alerted him to incidences of HIV-contaminated blood. In an election year, many observers feel that he is reacting strongly to the purported health risk from BSE in order



to promote his own credibility and that of his party, the Christian Democrats.

So far, Seehofer has been unable to convince the EC of the merits of his case. The commission is backing the claims of the British government that the risks from British beef are negligible.

Alison Abbott

Delayed pay rises still leave concerns in Hungary

Budapest. Scientists in Hungary have at last received a 30 per cent salary increase promised by the government in a law approved two years ago covering working conditions for all civil servants.

But concern remains at the way in which the law provides guaranteed positions to all scientists in government-funded research institutes, at a time when the Hungarian Academy of Sciences is trying to introduce more flexible employment practices.

The law was passed in 1992, but only took effect this year. Last autumn, concerned at the costs of a substantial increase in the salaries of all public employees, the government tried to renegotiate the new pay scales. It backed down after a protest from civil servants, including health workers, teachers and research scientists.

But it seems that the increases will cover only minimum wages, not the additional bonuses promised to scientists. Bonuses are paid to individuals who, for example, speak foreign languages or handle dangerous chemicals. They are often vital to scientists whose monthly wages average only US\$300.

Even though the issue of increased pay has been resolved, the more difficult problem remains of how the scientific community should handle the government's promise to provide secure posts for all researchers working for the Academy of Sciences.

Pál Venetianer, director of the academy's Biological Research Institute in Szeged, southern Hungary, says that short-term positions as postdoctoral research workers must be provided by academy institutes if a flexible system of working is to be developed.

Alison Abbott

Biotechnology alliances set to double

London. The number of strategic alliances between European biotechnology companies is likely to more than double over the next five years, according to a survey published this week by the consultants Ernst & Young Europe.

Unveiling the results of the company's first survey of the European biotechnology scene, *Europe Biotech 94: A new industry emerges*, Pieter Lucas, a consultant to the company, revealed that Europe has 386 biotechnology start-up companies, with the United Kingdom accounting for a third of them. All are building up their strategies to meet growth in the global marketplace, said Lucas. "CEOs [chief executive officers] expect alliances with pharmaceutical companies to increase by 50 per cent over the next five years. Alliances with European agricultural companies are expected to remain constant, while alliances with US ag-

ricultural companies are expected to increase threefold."

He added that CEOs are placing strong emphasis on receiving advance fees and royalty percentages from alliance partners. Even more important, however, is the structuring of collaborative research and development arrangements.

The survey showed that there is virtually a consensus on near-term strategic goals. The priorities are the acquisition of new products (40 per cent of respondents) and the strengthening of alliances with universities (38 per cent of companies surveyed).

Europe's biotechnology sector is already thinking about second-generation products, with over a third saying they are looking for opportunities to license in new technology. A similar number say that licensing rights to specific applications are high on the agenda.

Mike Ward

High-energy physics faces fight for survival

Washington. High-energy physics facilities in the United States face deep cuts in their operating budgets for the third year running, according to a number of laboratory chiefs who addressed the Department of Energy's High Energy Physics Advisory Panel (HEPAP) in Washington last week.

Members of the panel, which advises the government on its expenditure on the field, point out that a continuation of present trends would see these operating budgets reach zero by the year 2005. "What I'm worried about in this field is its survival," said Donald Hartill of Cornell University in New York state.

Fellow panellist Pierre Ramond of the University of Florida applied some dry humour to the situation: "In this society, if something is not healthy, you shoot it". HEPAP has heard this sort of doom-mongering before (*Nature* 362, 581; 1993); but this time, in the wake of last October's demise of the Superconducting Super Collider (SSC), the physicists may not be crying wolf.

Certainly any hopes that established high-energy physics facilities would benefit from the collapse of the mammoth Texan project have already evaporated. The Department of Energy is to spend \$640 million this year on closing the SSC, and \$180 million next year; but as the albatross falls away, the budget request for the rest of the programme remains stagnant at \$620 million.

Furthermore, because of new construction projects at Fermilab, Illinois, and the Stanford Linear Accelerator Center (SLAC) in California, operating budgets will dip by four per cent, for the third year running.

Fermilab, SLAC and the third main facility, at Brookhaven, New York, are each planning staff cuts to prevent further reductions in the proportion of the year for which they can operate their accelerators. Cash-strapped high-energy physicists are also worried that they will get saddled with run-

The disappearance of \$44,000 worth of personal computers from the site of the Superconducting Super Collider (SSC) at Waxahachie, Texas (right), in January has prompted a harsh security clampdown, according to SSC director John Peoples. At the same time, tensions between the Department of Energy (DoE) and the State of Texas are souring the atmosphere under which the project is being wound down.

The Federal Bureau of Investigation is looking into the computer loss, and the energy department, concerned about attracting the attention of congressmen such as John Dingell (Democrat, Michigan), has quadrupled the security budget on the site to \$265,000 a month. Peoples, who remains director of Fermilab but is spending three days a week at Waxahachie, told a meeting last week of the DoE's High Energy Physics Advisory Panel that the clampdown makes a difficult job harder.

Relations between University Research Associates — the SSC contractor — the DoE and the state of Texas are strained. But Peoples appears determined to remain cheerful in the face of adversity. "People are working very hard. There are still some



old wounds but we're doing our best," he says.

The main problem at SSC is how to retain people to carry out key tasks. Staff numbers at the laboratory have already fallen from 2,000 last September to 850, and will wind down to zero by September 1995. But the management issues which Peoples is responsible for will, he hopes, be resolved long before then. "I could be out of here by this summer," he says. □

ning yet another facility in Texas when the energy department and the state government strike a deal on what to do with the SSC site (see box). "The SSC is worse than irrelevant," says panel member Stewart Smith of Princeton University, New Jersey. "It's becoming a tax on the whole programme."

But these straitened circumstances also raise questions about the efficacy of HEPAP, an open forum born in the heady 1960s but perhaps ill-suited to help manage a programme in contraction. Faced with competing demands on diminishing resources, HEPAP appears unable to set clear priorities, because, if it does, it risks losing those

projects placed lower on the priority list, as well as the money that comes with them.

At the moment, a HEPAP sub-panel chaired by Sidney Drell of SLAC is preparing a report on how the programme should proceed post-SSC, and in particular on prospects for collaboration with European high-energy physics (*Nature* 367, 397; 1994). After that, HEPAP's chairman, Stan Wojcicki of Stanford University, would like another sub-panel to meet in private and set priorities for the rest of the programme. But several panel members warned that such a process would be fraught with danger.

In the meantime, Wojcicki is planning to write yet another letter to Hazel O'Leary, the Secretary of Energy, complaining about the budget allocation. But he is painfully aware of its likely ineffectiveness. "I'd like to do something more dramatic or effective, but I don't know what," he says.

"HEPAP is in the mood that the sun is setting," observes Bernard Hildebrand, a former head of physics research in the energy department. He suggests that the community should organize a small group of its best people to go out and sell what they are doing to key officials in Washington.

These would include officials in the energy department, the National Science Foundation, the White House science office and both houses of Congress. But the trouble is that Washington knows their faces; and the last time they came to lunch, they were selling an \$11-billion project that turned out to be a white elephant. **Colin Macilwain**

Academic aid sets out for Yugoslavia

Paris. A convoy of lorries will set out from Strasbourg in France next month to deliver 30 tonnes of books and academic journals to the universities of Zagreb and Belgrade in the former Yugoslavia.

The convoy, and an air shipment to Sarajevo, where the university library has been destroyed, is one of several projects organized by academics and students at universities in Strasbourg. The group, "Students and teachers for Sarajevo", chose Sarajevo as a focus for attention, but is working for peace throughout former Yugoslavia.

In particular, the group has invited 30 students from Bosnia-Herzegovina, Croatia, Serbia, Slovenia and Macedonia to take part

in a public conference in Strasbourg from 2 to 6 May. The conference, timed to coincide with a sitting of the European Parliament, will consider how universities can help in rebuilding peace, as well as wider issues raised by the Yugoslav war.

"We are trying to re-establish the bridges that have been broken between them", says Ronald Hancock, a researcher at the Institute of Biological Chemistry in Strasbourg, who is on sabbatical leave from Canada. For further information, contact "Etudiants-Enseignants pour Sarajevo", 4 rue Thiergarten, 67000 Strasbourg, France (Tel.: 33.88.75.56.58, Fax: 33.88.37.97.25).

Declan Butler

Health reforms 'could hit biotech companies'

Washington. The spectre of price controls as outlined by President Bill Clinton in his health-care reform package is already having a negative impact on the ability of US biotechnology companies to raise investment capital, according to a survey of such companies released in Washington last week.

The survey found that, if Clinton's pricing proposals are implemented unchanged, and if the US capital markets dry up as a result, most companies say that they will be forced to seek foreign partners and buyers, and to scale back their expenditure on research and development.

The survey, entitled *Price Controls and the Future of Biotechnology*, was undertaken by Robert M. Goldberg, a senior research fellow at the Gordon Public Policy Center at Brandeis University, who surveyed more than 200 companies involved in therapeutic drug development.

Replies were received from 107 companies, of which 73 were publicly owned. More than two-thirds of the companies were involved in developing drugs for the treatment of cancer, a half for treatment of AIDS and a half for infectious diseases.

Of the 81 per cent of the companies surveyed that planned to raise capital in 1993, 60 per cent said they either withdrew

or postponed stock offerings, or fell short of their funding target; 27 per cent met target and 13 per cent exceeded target.

The Clinton plan is not explicit about its proposed controls on drug pricing, but it contains two key provisions. One would be the creation of an advisory council on breakthrough therapeutic drugs which, along with the secretary of Health and Human Services (HHS), would review the "reasonableness" of a manufacturer's price for such drugs.

In addition, the secretary would be given authority to negotiate special rebates or discounts on all new drugs sold to Medicare patients (senior citizens). In cases where agreement cannot be reached, the secretary would have the power to blacklist the drug.

Industry officials and stock market analysts argue that, if implemented, such heavy-handed regulation — which they say virtually amounts to price controls — would scare off nervous investors, and drive investment away from what is already a high-risk sector.

They argue that, as it stands, the plan's definition of "reasonableness" is too narrow. By focusing heavily on unit manufacturing costs, it fails to take into account the drug failures within a company, as well as economic benefits, such as reduced costs

associated with less surgical intervention and shorter hospital stays.

"The board doesn't bother me as much as the reimbursement exclusion power of HHS," says Jeffrey Casdin, senior vice-president for biotechnology investment research at Oppenheimer & Co. "If the system can't allow for a huge winner every once in a while, it's very hard to convince people to put up capital knowing they're going to lose most of the time."

Stelios Papadopoulos, managing director of investment banking at PaineWebber, points out that, faced with possible caps on the prices of new drugs, people will simply shift their investment opportunities elsewhere. Papadopoulos compares investing in biotechnology companies to buying lottery tickets: people buy into them in the hope that if they win, the winnings will change their lives. "Why would anybody buy a lottery ticket if the maximum [payout] was \$1,000, even if the odds were great?"

According to Goldberg's survey, the prospects facing biotechnology companies for raising capital this year appear no more than lukewarm. Only 55 per cent of companies rate their chances as fair, while 23 per cent say they are discouraged at the prospect.

Diane Gershon

French fossil-hunters defend the value of public support

Paris. A leading French palaeontologist last week defended the value of amateur fossil hunters to research workers, amid renewed calls for the introduction of legislation to protect sites in France from collectors.

The lack of legislation had been highlighted by the decision of Jean-Pierre Maggi, the mayor of the small town of Velaux near Marseille, to bring proceedings against a fossil hunter, Xavier Valentin, for "spoiling the image of the town".

Valentin had removed dinosaur bones from a nearby site, and Maggi says that such sites should be protected for "future generations who will have more scientific means available". But Eric Buffetaut, a Centre National de la Recherche Scientifique (CNRS) researcher at the Laboratory of Palaeontology and Stratigraphy at the University of Paris VI, says the mayor's argument reflects widespread misunderstanding of the importance of amateur palaeontologists to research.

According to Buffetaut, researchers lack the resources to dig the many areas where fossils occur rarely. In particular, he says, amateurs often find recently exposed fossils in cliffs or sites uncovered by floods, which would otherwise be destroyed by erosion.

Buffetaut argues that amateurs such as Valentin have contributed many valuable specimens to his group, and to the museum

at Esperaza (see below). Although commercial or poorly trained fossil hunters are a big problem, any legislation should be restricted to protecting geologically stable sites of special scientific interest, he says.

Buffetaut has also been involved in a

graduate, confirmed that the area was the richest site for dinosaur remains in France.

Le Loeuff planned to use income from the museum to support his research. Few palaeontologists are being recruited in France, and although Le Loeuff is applying for a post in the CNRS, it has twice rejected him.

But according to Buffetaut, who is president of the non-profit association set up to run the museum, the council kept three-quarters of the income, instead of just collecting rent as agreed. Buffetaut also says that the council refused to hand over the running of the museum, as it had promised to do. This prevented the association from seeking other partners to help expand the museum, and caused administrative absurdities; researchers, for example, could not enter the museum after 6 p.m.

Last November, Le Loeuff decided to protest. He hid the exhibits and closed the museum, which had attracted more than 40,000 visitors in the year after it opened in 1992. As a result, the number of visitors to the town plunged, and last month the council backed down. The association will now run the museum, and "we will get all the money", says Le Loeuff.

Declan Butler

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Esperaza Museum: open again to visitors.

separate dispute that had threatened to close a popular dinosaur museum in Esperaza in southern France. The dispute ended last month when its creator, palaeontologist Jean Le Loeuff, won a battle with the local council over who should control the museum.

Le Loeuff first went to Esperaza in 1988, shortly after the discovery there of fossil dinosaur eggs. Subsequent digs by Buffetaut's laboratory in Paris, where he was a post-

Philippe Rielly/Eurelia/SPL

Japan seeks to reverse drop in interest in science

Tokyo. Faced with a declining interest among young Japanese in pursuing careers in science and engineering, Japan's Ministry of Education, Science and Culture last week set up an advisory committee made up largely of young researchers to produce ideas about ways of reversing this trend.

The tendency for young Japanese to turn their backs on science and engineering has been evident for several years. In 1989, for example, the National Institute of Science and Technology Policy (NISTEP) reported a sudden drop in the number of science and engineering graduates entering manufacturing industry, and a surge in those going into finance (*Nature* 338, 102; 1989).

At the time, it was not clear whether this was a temporary phenomenon associated with Japan's then booming stock market. More recent studies suggest the change is deep-rooted. For example, the annual white paper (policy document) on science and technology, released by the Science and Technology Agency in December, produces several indicators to show young people's declining interest in science.

In the early 1980s, more than half (56 per cent) of young people in their 20s expressed interest in news and topics concerning science and technology, larger than the percentage for older age groups. But by November 1991, the proportion of 20-year-olds had fallen to 41 per cent, far less than in older groups.

This trend is reflected in the percentage of applicants to science and engineering faculties at universities, which, according to a ministry of education survey, peaked at more than 25 per cent in 1986, but has since fallen to less than 20 per cent.

Furthermore, another survey last month by the education ministry of 57 deans of engineering and science faculties at national and private universities shows that by far the majority (44) feel that the best students are no longer choosing to study science and engineering.

In a departure from tradition, the ministry's committee, which will be chaired by Yasuharu Suematsu, emeritus professor of Tokyo Institute of Technology, includes many relatively young members still in their 30s and 40s, and three women.

The committee is expected to make its recommendations in June and is likely to call for dramatic improvements in the research environment in universities, for revisions in curricula, and for more open days for schoolchildren to visit university laboratories. It is also expected to urge universities and industry to form a closer alliance to improve career prospects for scientists and engineers.

David Swinbanks

Soros pledges more money to back Russian science

Moscow. George Soros, the Hungarian born financier, has agreed to provide a further \$12.5 million to the International Science Foundation (ISF), the private organization which he set up two years ago with an initial grant of \$100 million to provide support for Russian scientists.

Soros' decision, which was announced in Moscow on Monday, follows a promise by the Russian government that it is prepared to contribute a similar sum to the foundation. Earlier, Soros had said that he would only provide additional funding for the ISF if the organization was able to raise matching funds from other sources.

Speaking at a press conference which was also attended by Boris Saltykov, the Russian minister for science and technology policy, and James Watson, the Nobel laureate and current chairman of the ISF, Soros said that he had been impressed by the work of the foundation over its first two years, and in particular the efficiency with which its activities had been organized.

He also emphasized his desire to continue to provide support for other activities of the foundation, including the assistance provided to scientists for the purchase of foreign journals and, most importantly, for the development of electronic communications.

During his visit to Moscow, Soros signed an agreement with the Russian government aimed at developing the use of computer networks in Russia. He expressed a desire to see the new networks enlarged to cover

libraries and other cultural organizations.

In addition to the new financial commitment from both Soros and the Russian government, the foundation has set itself the target of raising \$60 million from both the United States and other foreign governments. At present, however, there is little indication of how much of this money will be made available.

The results of the earlier round of awards made by the ISF show that about one fifth of the nine thousand applications received were approved, with the size of most of the grants varying from \$10,000 to \$50,000.

According to Pavel Arseniev, the head of the ISF programme, a lack of funds meant that the value of the largest grant had to be cut from a promised \$100,000 to \$69,000. But he says that the overall success of the programme is reflected in the large number of applicants for awards.

Despite the lack of funds, the ISF staff decided to increase the number of grants to be awarded, reducing the payment schedule from two to one-and-a-half years. As a result, financing will end in the summer of 1995, rather than in December as originally planned.

Meeting in Moscow on Monday, the executive board of the ISF approved recommendation from its expert committees for the distribution of about \$32 million in long-term grants. A further \$13 million is due to be awarded in the second round this autumn.

Carl Levitin

Marie Curie 'should rest in Panthéon'

Paris. Marie Curie may become the first woman to be buried in France's national mausoleum, the Panthéon. The proposal was made on international women's day last week by François Mitterrand, the French president, in response to protests from leading female personalities—including Simone Veil the health minister—that at the end of the twentieth century it was absurd that the official recognition symbolized by the Panthéon was still reserved for men.

In 1903, Curie became the first woman to win the Nobel Prize for Physics, along with her husband Pierre, for the discovery of radioactivity.

In 1911, she became the first woman to win a second Nobel Prize, this time for chemistry. But in the same year the

French Academy of Sciences turned down her application for entry, saying "let in anyone, except women". Her daughter Irene, won the Nobel Prize for chemistry in 1935, with her husband Frédéric.

Technically speaking, Marie Curie will be the second woman to be buried with figures such as Voltaire, Rousseau, and Victor Hugo. The remains of Marguerite Berthelot were placed beside those of her husband, the chemist Marcellin, as he had stipulated in his will.

All that now stands in the way of Mitterrand's wish is the consent of Marie-Curie's relatives. If they agree, they are likely to ask that Pierre accompany her; and the Panthéon will have to insert "*femmes*" into its motto, "*aux grands hommes*."

Declan Butler



Belated recognition.

Employment in a rainforest

SIR — Are any readers interested in working in a tropical rainforest?

I recently paid a visit on behalf of the Royal Society to the Danum Valley Field Centre in the state of Sabah, East Malaysia (formerly British North Borneo). Nearly one million hectares of rainforest in Sabah (about 15 per cent of the total land mass) have been allocated to the Sabah Foundation, a non-profit-making trust set up by the local people to support a programme of education, welfare and health. Two large areas remain untouched and committed as conservation areas; most of the rest is used for 'sustainable forestry'. Initially, in 1966, this involved logging followed by regeneration so that the productive timber resources were maintained in perpetuity. More recently the foundation has come to realize that future generations in Sabah and elsewhere require 'sustainability' not just of the timber resources but of the entire ecosystem. They are now seeking to adopt forestry practices that protect the environment and do minimal damage to the tremendous diversity of plant and animal life that characterizes tropical rainforests.

The field centre sits at the boundary between untouched primary forest in the 438-km² Danum Valley Conservation Area and forest that has been logged at various known times by various specified procedures. It offers unique opportunities to study the natural biodiversity of the forest, how it is disturbed by logging and to what extent it regenerates. Research projects are in progress on rainforest dynamics, soil water and stream hydrology, erosion, rehabilitation by enrichment planting, insect diversity, fish populations, mouse deer radiotracking, orang-utan feeding habits and much more. The Royal Society provides some core support and scientific coordination and the centre is managed by a consortium that includes the Sabah Foundation, the Sabah Forestry Development, the local Ministry of Tourism and Development and the Sabah Campus of the National University of Malaysia.

The research opportunities are unbounded, and the resources of the centre more than adequate for the number of researchers working there. Malaysian scientists involved with the programme are enthusiastic but few in number and inevitably much of their time is taken up by administration and teaching the next generation of science students. British scientists are welcome: either senior scientists for shorter working visits, if possible combined with lecture courses at the university, or postdocs or graduate students working for higher degrees.

For further information, write to either Dr Clive Marsh, Assistant General Mana-

ger (Conservation and Environmental Services), Forestry Division, Innoprise Corporation Sdn Bhd, PO Box 11623, 88817 Kota Kinabalu, Sabah, Malaysia, or Dr Stephen Sutton, Royal Society Research Programme Coordinator, Heigh Head, Mewith, Bentham, Lancashire LA2 7AU, UK.

Anne McLaren
(Foreign Secretary)
Royal Society,
6 Carlton House Terrace,
London SW1Y 5AG, UK

Education in India

SIR — I should like to suggest some ways of remedying the present parlous state of higher education in India.

In the first place, science and technology should be included with the education portfolio at the centre, as it was when Maulana Azad was Union Education Minister; when Justice M. C. Chagla was made education minister in Nehru's cabinet, science and technology were included in his portfolio.

Second, the universities should be removed from the control of state governments and brought under the direct administrative control of central government, where the University Grants Commission (UGC) can monitor academic growth properly rather than merely distributing funds. This will help to create a national outlook on higher education.

Third, the UGC budget, which has remained constant during the past three years, should be increased in the light of current needs.

Fourth, 10 per cent of the science and technology budget should be earmarked for training in universities and colleges. The present allocation to science of 0.9 per cent of gross national product (GNP) is even less than what was spent in the 1980s (1.3 per cent). This needs to be raised to a minimum of 8 per cent of GNP.

What has gone wrong with Indian universities? When the British left India in 1947, they left four universities — Allahabad, Bombay, Calcutta and Madras — with high standards of scholarship and learning and with full autonomy. When India became independent, universities came under the control of state governments and, in the name of economic control, universities lost their autonomy and petty bureaucrats dictated day-to-day functioning. The downward path was also heralded with the creation of Institutes of Technology (IITs) totally divorced from the academic life of universities. Universities were starved of funds as major resources were diverted to the creation of IITs. With the passage of time, the IITs

have become academic ciphers like the universities. There is talk in India of granting national laboratories university status and allowing them to award degrees. This would drive the last nails into the coffin of the already moribund university system.

Finally, India must realize that national laboratories must be made subservient to industrial research needs. National laboratories must help towards industrial development to make India self-sufficient in production and in technical know-how. Basic science and research should, by and large, be left to the universities. Of course, no laboratory can function without some work in basic sciences, but this should be seen as a secondary function of the laboratories, and not their primary function.

John Maddox has rightly said (*Nature* 366, 627; 1993) that "India is lucky in knowing what it needs". The knowledge of one's needs is not sufficient if one does not have the moral courage and mental calibre to muster means. The demoralized academic community in India needs the means to steer its path through a self-defeating bureaucratic system coupled with political ignorance of the importance of academic and scientific excellence for attaining social justice.

A. N. Malviya
5, rue Blaise Pascal,
67084 Strasbourg, France

Jobs for the girls?

SIR — The discussion of "social science and the new world order" (*Nature* 366, 403; 1993) raises a pertinent question — can the example of Kerala, where the birthrate has fallen dramatically, be copied elsewhere? Even other Indian states could not emulate Kerala. There is thus a need to look more critically at this success story.

Even if it is possible to provide the necessary infrastructure for education, there is no guarantee that it would be used, because there has to be a perceived benefit from education that would motivate families to educate their children (especially girls).

Kerala succeeded by providing a perceived link between female literacy and female job opportunities. In Kerala, women make up a much higher proportion of the workforce than elsewhere in India — 35 per cent overall, 45 per cent in the private sector. Education without job opportunities will not bring about the classical demographic transition.

Shashi Kant
Department of Medicine,
L. R. Murmu
Department of Surgery,
All India Institute of Medical Sciences,
New Delhi 110 029, India

HIV in Zambia : myth or monster?

P. Godfrey-Faussett, R. Baggailey, G. Scott & M. Sichone

Despite attempts by a few journalists to portray the spread of AIDS in African countries as fictitious, the situation in Zambia shows that such views are completely misguided.

TEN years ago, Anne Bayley, working at the University Teaching Hospital in Lusaka, Zambia, noticed that not only was the number of her patients with Kaposi's sarcoma increasing, but the characteristics of the disease had changed¹. Previously, patients with this condition had been older men with the disease predominantly affecting the legs and responding satisfactorily to treatment. The new patients were often young women, with widespread lesions on the trunk, in the lymph nodes, in the mouth and in the lungs. Treatment was much less satisfactory: patients were dying with generalized oedema. Over the next few years, the number of atypical cases continued to increase and was shown to be associated with antibodies against HTLV-III (as HIV-1 was then known)². Any physician or surgeon working in the 1970s and 1980s would agree that a new pattern of disease had emerged. AIDS has not been present for decades in Africa, unsuspected until clinicians in the United States drew attention to it.

In contrast to the initial studies in the United States, serological surveys in Zambia³ and elsewhere in Africa⁴ showed that HIV seropositivity was equally distributed between the sexes. The highest seropositivity occurred in young adults, with the peak in women occurring at a younger age than that in men. Seropositivity was strongly associated with the number of reported sexual partners and previous episodes of sexually transmitted diseases. Since then, repeated surveys have demonstrated rising rates in the general population but seropositivity is still associated with the sexually active age group, their infants or those who have received blood transfusions.

During the past decade, the attitudes both of individuals and of government institutions has changed from prophecies of apocalypse to a more measured evaluation of risk. This shift of emphasis has been greeted with enthusiasm by those who believe that HIV is an irrelevance in AIDS, and seems to have encouraged a new revisionism that has found expression in the British⁵ and the African^{6,7} press. The theme of these stories is that AIDS in Africa is a myth; that longstanding diseases such as diarrhoea and tuberculosis represent more immediate problems; and that malnutrition and poverty, together with climate and political disruption, account for the worsening health of people observed in the past 10 years. Pro-

ponents of this view argue that the myth is being propagated to sustain the resources of those researching AIDS at the expense of other causes of morbidity and, more subversively, that this myth forms part of a conspiracy to prevent development and investment in Africa.

Few people who live here could believe these stories. Deaths in early adulthood have become an accepted part of Lusakan life. In surveys carried out among 767 university students, 397 employees of companies and 500 of those attending voluntary HIV counselling and testing centres, 39, 23 and 29 %, respectively, of those interviewed had lost relatives to what they believed to be AIDS and a further 12, 8.4 and 11% had a relative known to be HIV positive⁸.

Zambia is a landlocked country in southern Africa with a population of about 8 million and an average growth rate of 3.2%. Like most sub-saharan countries, dependency is very high: with 48% of the population under the age of 15 and 3% over 65, there is only 1 working-age adult for each dependent child. The weak economy, huge debts and poor health infrastructure have been compounded by the recent drought: 19.1% of children die before their fifth birthday.

Despite being a predominantly Christian society, traditional beliefs are still widely held, nowhere more so than about health. HIV and AIDS still carry a powerful social stigma — less than half the people who come for voluntary, confidential HIV tests are able to discuss their test with their partner, whatever the result.

More than 75% of people consult a traditional healer before attending a clinic or hospital, and deaths of people who have had AIDS-like symptoms are often ascribed to witchcraft. Rather than cause distress in the families of the deceased, doctors rarely mention HIV or AIDS on death certificates. In many areas, forms for the notification of AIDS to the Ministry of Health are not available and resources are inadequate to perform HIV tests. Even so, 7,761 cases of AIDS and a further 21,973 cases of AIDS-related complex had been reported to the National AIDS Prevention and Control Programme by October 1993. Dr Roland Msiska, the head of that programme, is confident that these figures are gross underestimates.

Given that a positive HIV test cannot be included in the diagnostic criteria for AIDS in Zambia, and the ineffectiveness of the reporting system, can we have any certainty that HIV is the monster that some have led us to believe? Has the HIV industry become self-sustaining, propagating (knowingly or not) dogma about an epidemic that does not exist? There are two main arguments that convince us that there is indeed a monster threatening those whom the country can least afford to lose: the serological surveys in different groups; and the mortality data.

Since the first tests were developed, Zambia has reported alarming seroprevalences of anti-HIV antibodies. Sequential surveys among women attending antenatal clinics in various rural and urban sites throughout the country have shown



A growing reality of Zambian life — AIDS education poster in Lusaka.

rising rates with 33% of such women in the high-density areas of Lusaka being infected (Ministry of Health figures, 1992).

Much has been made of the issue of the accuracy of tests for anti-HIV antibodies used in African countries. Initial observations were made using crude antigen preparations from HTLV-III, which caused many false-positive reactions⁹. Subsequent test systems using either recombinant proteins or competitive inhibition are much more specific¹⁰. But even a specificity of 99.5% is disastrous if a test is to be used for surveillance in populations where there is not much disease. If the prevalence is 1 in 1,000 (0.1%), for example, there will be 5 false positives for each true positive recorded, or a measured seroprevalence of 0.6%. But if the same test is used for surveillance in a population with a 25% prevalence of disease, the measured seroprevalence will be 25.1%. Even using a test with a specificity as poor as 95%, the measured seroprevalence would be 28.5% (assuming a sensitivity of 99%). Thus, concern with the specificity of tests is highly appropriate in low-prevalence areas but matters less when the prevalence is high: the distinction between 25 and 28.5% destroys few myths.

What distinguishes those who are seropositive for HIV from those who are seronegative? The most striking thing both in Zambia and elsewhere in Africa is that the former are more likely to die^{11,12}. Atypical Kaposi's sarcoma, tuberculosis and some bacterial, fungal and protozoal diseases are commoner in those who are seropositive, as are chronic diarrhoea, wasting and lymphadenopathy. So HIV appears to be a marker of an underlying immunodeficiency. The suggestion that the cause of this immunodeficiency could be malnutrition or intravenous drug use does not stand up to scrutiny. HIV seropositivity correlates with years in full-time education³; intravenous drug use is not widespread in Lusaka.

In the past decade, tuberculosis notification rates in Zambia have climbed from 120 to 310 per 100,000 population per year (Ministry of Health figures 1992). In 1989, a cohort of patients was recruited and followed at the Chest Clinic in Lusaka¹³. Of these, 73% were HIV seropositive. This study and those from other African states showed major differences in the course of tuberculosis, the reaction to drugs and the outcome of treatment in those patients who have a positive HIV test (whatever that may signify). Such individuals are up to six times as likely to die within a year of starting treatment^{14,15}; up to six times as likely to develop dangerous skin reactions to thiazetazone (an anti-tuberculosis drug widely used in Africa)^{15,16}; and up to thirty times more likely to develop recurrent disease¹⁷ after completing treatment as patients whose HIV test is negative. Nor is this because

those with a positive HIV test were the most impoverished — on the contrary, they tended to be better educated and to live in less crowded conditions than those with negative tests¹⁸.

There is more and more HIV infection in the community. But if those who are HIV seropositive are more likely to die, where are all the deaths? Cohort studies in Britain and the United States show median survivals of more than a decade from the time of infection. Similar data are very sparse from Africa, and nationwide mortality statistics are difficult to interpret because of delays and incomplete records. But because Zambian employers traditionally offer considerable financial assistance to the families of employees who die, mortality figures for Zambian industry are accurate. A survey of 33 businesses in Lusaka and the Copperbelt, which included a workforce of more than 10,000, shows that from 1987 to 1993 mortality rose from 0.25 to 1.83% per annum. Similarly, a survey of nurses employed in two hospitals in the southern province shows mortality rising from 0.2 to 2.7% per annum¹⁹. These are not the starving and destitute who are dying; they are people with an income who are contributing to the economy. We remain convinced that there is a disaster unfolding in Africa in which the most productive members of society are being destroyed. HIV antibodies predict those who will die, but most deaths are among people who have not been tested and so are unlikely to be recorded as HIV-related.

The natural history of HIV infection in Africa remains obscure. One thing that is clear, however, is that in Zambia, as elsewhere, emotions, drives and addictions sometimes overwhelm the rational objective of self-preservation. Everyone is very aware of the risk of catching AIDS from sexual partners, and everyone has been told a thousand times that becoming less promiscuous and using condoms are the only known strategies to reduce their risk of becoming HIV positive. But maintaining a cool rationality 24 hours a day, 365 days a year, is beyond many people's natural abilities. Few people invariably translate knowledge into behaviour. Even counsellors who deal with AIDS-related tragedies every day have been known to abandon their own advice (and free condoms) after hours. It is not in the cold light of day that the *Sunday Times*-style revisionism is most dangerous — Zambians have seen too much AIDS to believe it. But by night it could become one factor in an individual's or couple's decision to abandon caution.

There is considerable ambivalence in Zambia about open discussion of people's sexual lives and of condoms. Christian moralists argue that fidelity should render condoms irrelevant, and Zambians are reluctant to use them. But the fight is on.

The new Minister of Health, Michael Sata, a colourful and influential man of the people, advocates a more direct approach. On a popular television discussion programme last month he warned of the dangers of ignoring the issues and challenged doctors to be more open about the diagnosis of AIDS and to record the cause of death more precisely.

Some of us have been concerned with the AIDS epidemic in Zambia for 10 years. The hope that we will wake up one morning to find that the whole thing has been a ghastly nightmare is a familiar one. But the loss of friends, and friends of friends, makes it all too real. It is one thing to be 6,000 miles from the killing ground, peering through a badly focused telescope, but quite another to be here.

The AIDS-is-a-myth lobby is fond of portraying itself as a courageous David taking on the Goliath of scientific orthodoxy. However, its revisionism is hardly a deeply insightful, wild surmise like de Broglie's suggestion that particles are waves. Nor is it on a par with Sheldrake's amusing, cheeky challenge to orthodoxy (and *Nature* in particular) in the shape of morphic resonance. The AIDS-is-a-myth myth flies in the face of every explanation of current events which could conceivably be true, given the easily accessible facts. And it is downright dangerous. The awareness-raising that has been going on in Zambia, and in other countries, for a decade will not pay dividends in the form of safer sexual behaviour if people are offered a reason to ignore an unpleasant truth. □

P. Godfrey-Faussett is in the ZAMBART project, Department of Medicine, University Teaching Hospital, PO Box 50110, Lusaka, Zambia and is also at the Department of Clinical Sciences, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. R. Baggeley is at the Kara Counselling and Training Trust, Lusaka and at the London School of Hygiene and Tropical Medicine. G. Scott is Member of Parliament for Mpika, Northern Province, Zambia. M. Sichone is manager of the National Tuberculosis Control Programme, Zambian Ministry of Health.

1. Bayley, A. C. *Lancet* **I**, 1318–1320 (1984).
2. Bayley, A. C. *et al.* *Lancet* **I**, 359–361 (1985).
3. Melbye, M. B. *et al.* *Lancet* **II**, 1113–1115 (1986).
4. Nkowane, B. M. *AIDS* **5** (suppl 1), S7–S15 (1991).
5. Hodgkinson, N. *Sunday Times* 10–11 (3 Oct 1993).
6. Versi, A. *New African* 8–11 (Dec. 1993).
7. Robbins, D. *The Star, Johannesburg* 19 (27 Jan 1994).
8. Abstract WPE 130 Marakesh VIII Int. Conf. AIDS Africa.
9. Biggar, R. J. *et al.* *Lancet* **II**, 520–523 (1985).
10. WHO Global Programme on AIDS WHO Wkly Epidem. Rec. **67**, 145–149 (1992).
11. Lucas, S. B. *et al.* *AIDS* **7**, 1569–1579 (1993).
12. Gilks, C. F. *et al.* *Q. J. Med.* **82**, 25–32 (1992).
13. Elliott, A. M. *et al.* *J. Trop. Med. Hyg.* **96**, 1–11 (1993).
14. Nunn, P. *et al.* *Am. Rev. Respir. Dis.* **146**, 849 (1992).
15. Perriens, J. H., Colebunders, R. L. & Karahunga, C. *Am. Rev. Respir. Dis.* **144**, 750–755 (1991).
16. Nunn, P. *et al.* *Lancet* **337**, 627–630 (1991).
17. Hawken, M. *et al.* *Lancet* **342**, 332–337 (1993).
18. Elliott, A. M. *et al.* *Br. Med. J.* **301**, 412–415 (1990).
19. Buve, A. *et al.* *AIDS* (in the press).

Defending science against anti-science

There is a need for concerted action against the forces of anti-science; simply ignoring the critics will not suffice to counter the phenomenon.

JOHN Ziman's recent review of Gerald Holton's latest book (*Nature* 367, 522; 1994) will have prompted many people to read not just the review, with its stirring warnings of the dangers of "anti-science" (part of Holton's title), but the book itself. That is what publishers and authors hope that reviewers will do for them. Ziman has done the job excellently.

But many who follow Ziman's signpost may be disappointed. Holton's essay on "The anti-science phenomenon" is the last in a collection that begins with an essay on Ernst Mach's influence on science in the early decades of this century. Legend (and Einstein's slightly ambiguous expression of gratitude) has it that Mach's positivism helped in constructing the foundation of relativity. It remains in dispute whether Mach repudiated his delight in what his disciple had wrought. The truth is that Mach stimulated "the unity of science" movement that sporadically occupied European and American science in the 1920s and 1930s. In that sense, Mach is a good foil for anti-science.

The disappointment is not so much with substance as with style. Holton's defence of science against its detractors is, inevitably in the circumstances, cool, even quasi-philosophical. It is all there, but too measured for many tastes, and the present need.

Take, for example, astrology. It is a plain fact that astrology is a pack of lies in the literal sense; those who peddle horoscopes do so on an explicit set of statements about the real world that cannot be correct. There is no evidence that the positions of the planets can affect human behaviour, nor any plausible mechanism by which they could do so. It would not matter if the lies were told in some other context, say the alleged link between stock-exchange prices and the popularity of rock-and-roll music. That they are told, and believed by countless innocents, in flat contradiction of the more objective view of the world accumulated over centuries, means that each and every horoscope is, by denying the objective view of the planets, an attack on the probity of science.

Holton makes that point, and also notes that the scientific profession is docile in the face of what is really a torrent of attack. But is it not disgraceful that there should be such general and benign tolerance of astrology (and other mumbo-jumbo such as faith healing, water divining and spiritualism), apparently on the grounds that they are the harmless pursuits of people who are not scientists? Would other professionals, lawyers or accountants say, be as tolerant of

public belief that undermined the integrity of their work — and, potentially, their livelihood?

Religion is something else again. As the long (and soon to end) correspondence in *Nature* has shown, many professional scientists are deeply religious, often justifying their belief on the grounds that "science cannot know everything". That science does not know everything is, of course, accepted generally. Professional people might even more openly acknowledge that much of what is known may be incorrectly (or incompletely) known. The idea that religion may be a way of organizing one's appraisal of one's place in the world is not very different from what astrologers tell their clients. In other words, it may not be long before the practice of religion must be regarded as anti-science. Holton, who skirts round this issue, is nevertheless correct in saying that 'creation science', as its practitioners call it, deserves the fiercest counter-attack, especially now that so many of its practitioners have qualified as scientists and engineers, and so are all the more insidious because they know the language.

In reality, of course, there are many more contemporary attacks on science than astrology and religion. Moreover, they are more direct, and often sustained by people even better qualified in science or engineering than those recruited by the creationists. Yet that has not prevented the engagement of these well advised pressure groups in public issues by exploiting the willingness of the general public to be frightened by issues they do not understand. Similar problems arise with nuclear power and genetics; the pressure groups draw attention to many problems that need to be resolved, but then frighten the rest of the world into believing that problems not yet tackled spell potential catastrophe for some section of the community or the whole of it.

Because those who formulate these exaggerated arguments are often themselves members of the scientific community by qualification, they deserve respect, and not the disdain they usually enjoy. That is an error that creates more trouble than it saves. The usual reaction to those who perpetrate inferior contributions in the scientific literature, that their publications do not command serious attention, does not (or should not) apply to false statements of the view of the world equated with science. They deserve, but do not get, the treatment appropriate for astrology.

There is an obvious danger in a policy by

which the scientific community defends itself against anti-science exclusively by saying that its critics are mistaken. That would seem mean-minded to say the least of it. That is why Holton's references to Mach, and to Mach's earnest contemporaries, is an especially apt lesson for the present. Then, there was a 'world view' of science from which Mach's advocacy of the unity of science sprang. Now there are a half a dozen conflicting views (which Holton lists) which make it difficult for science to present a coherent counter to anti-science.

There are several distractions, of which the long argument about the death of physics is the most irrelevant. The issue goes back to the closing decades of the nineteenth century, when the practitioners of newtonian mechanics held that the problems they could not understand represented a failure of physics, but in the modern world is represented by those who argue that it needs only one last push to unify the four fundamental forces of nature, whereupon the whole of fundamental physics will be represented by a single formula. The optimism would be believable if anybody could explain how the missing step — the quantization of gravity — can be taken without changing our view of what the world is really like.

Yet Holton is right to insist that science needs a description of what it says (or will say) about the world that can lend coherence to its defence against its critics. For that matter, with this millennium coming to an end, it could be held that science needs such a description for its own belief in what it is about. These are the essential ingredients of what needs saying.

First, truth is not absolute, but rather a series of working hypotheses that are defined by increasingly rigorous rules; that means that the range of credible heterodoxy steadily diminishes. Second, to stifle arid arguments on questions such as the propriety of reductionism, it would suffice to say that science takes as its general working hypothesis the notion that all systems (in physics and biology) contain the ingredients and forces of their own evolution. That does not mean that the future is predictable from the present, for accidents (such as the K/T impact) will always happen. And finally, the purpose of it all is to understand the world more fully for no other purpose than the enhancement of the general enlightenment. Holton has done a public service in drawing attention to the threat of anti-science. He needs a polemicist to take the fight against it.

John Maddox

Beauty and the beholder

Nancy L. Etcoff

A BEAUTIFUL human face inspires pleasure and interest and often attracts riveted attention. But what constitutes beauty? There must be some general understanding of the concept, however vaguely defined; for instance, even two-month-old infants prefer to gaze at faces that adults find attractive¹. In recent years scientists have joined plastic surgeons and cosmetic companies in taking a deep interest in the question, and one such research group — Perrett, May and Yoshikawa — reports its latest results in this issue². The findings are scarcely definitive; in this area it is difficult to conceive of any that could be. But they add a new dimension to psychological, anthropological and biological thought on the subject.

The common notion has been that beauty is in the eye of the beholder, that individual attraction is not predictable beyond our knowledge of a person's particular culture, historical era or personal history. Beauty has also become a politicized issue. In her bestselling book, *The Beauty Myth*, Naomi Wolf³ argues that there is no such thing as a quality called beauty that "objectively and universally exists". In a society such as the present-day United States, where women earn consistently more than men in only two professions (modelling and prostitution), and where both cosmetics and weight loss programmes are enormously profitable industries, one can see the obvious concern in even contemplating perception of

beauty as a universal.

But the assumption that beauty is an arbitrary cultural convention may simply not be true. Perrett *et al.* belong to a growing body of scientists who are beginning to challenge it, just as scientists have begun to question anew many other assumptions about the relationship between human behaviour and culture. As Cosmides, Tooby and Barkow⁴ have pointed out: "Culture is not causeless and disembodied. It is generated in rich and intricate ways by information-processing mechanisms situated in human minds. These mechanisms are, in turn, the elaborately sculpted product of the evolutionary process. Therefore, to understand the relationship between biology and culture one must first understand the architecture of our evolved psychology."

Until the 1960s, it was believed that languages could vary arbitrarily and without limit, but now there is a consensus among linguists that there is a universal grammar underlying this diversity⁵. Similarly, it was once thought that facial expressions of emotion could vary arbitrarily across cultures, until Ekman and others showed that a wide variety of emotions are expressed cross-culturally by the same facial movements⁶. Ekman made the important distinction between the expression of emotion and the cultural variation that may exist in the rules for displaying those emotions. Likewise, although some aspects of judgements of

human facial beauty may be influenced by culture or individual history, the general geometric features of a face that give rise to perception of beauty may be universal, and the perception of these features may be governed by circuits shaped by natural selection in the human brain.

What is the evidence for this, and what would a universally beautiful face look like? Donald Symons, an anthropologist, proposed⁷ that beauty is averageness — the average values of the features of faces in a human population. Symons made the prediction on the basis of evolutionary biology and the principle that, during most periods, evolutionary pressures operate against the extremes of the population⁸. If this stabilizing selection principle is at work, and people with average physical properties have the best chance of survival, one would maximize fitness by being attracted to and mating with partners displaying such properties. There would thus be selection pressure to find average features attractive.

The hypothesis was tested in 1990 by Langlois and Roggman⁹, who used a computerized version of a technique developed by Galton a century earlier¹⁰ (see box). Galton superimposed photographs to create composites of faces, and to his surprise and frustration (one of his aims had been to create a prototypical criminal) the composite appeared more attractive than any of the individual photos that went into it. Langlois and Roggman confirmed this effect using ratings by college students of computer-generated composite faces.

Averageness, however, need not be the only criterion for beauty that natural selection might have favoured. When there is competition for partners — the precondition for Darwin's 'sexual selection' — those animals with certain kinds of extreme traits can often be preferred¹¹. Such extreme traits, the peacock's tail being the most famous example, can be a sign of the owner's innate resistance to disease and parasites, or an advertisement of its ability to gain sufficient resources to be able to 'afford' the flamboyant trait. Any disadvantage of the extremeness of the trait might be offset by the advantage of its attractiveness to potential mates.

Evolutionary biologists have thus painted two portraits of the face of beauty, one composed of features that are at the mean of the population, and another composed of at least some features at the population extreme. Perrett and colleagues attempt to discern whether averageness alone is beauty or whether we too may prefer the peacock. Using composites of either Caucasian or Japanese faces, they found that in both cases faces rated as 'attractive' were preferred to the composite of the sample from which the faces were selected; moreover, an attractive composite could be made more attractive

Galton and composite portraits.

In Francis Galton's pioneering paper of 1878, which Nancy Etcoff mentions in the main article, Galton describes how he "caused trials to be made" of ways of extracting typical characteristics from drawings of photographs of different faces. He reports how a photographic process, involving a stereoscope,

... enables us to obtain with mechanical precision a generalised picture; one that represents no man in particular, but portrays an imaginary figure, possessing the average features of any group of men. These ideal faces have a surprising air of reality. Nobody who glanced at one of them for the first time, would doubt its being the likeness of a living person.

Technologies change. The woodcut reproduced here appeared in the paper, but has been computer enhanced to improve the quality of reproduction.



It came with these comments:

... This composite is made out of only three components, and its three-fold origin is to be traced in the ears, and the buttons to the vest. To the credit of my judgment the original photograph is a very exact average of its components However the judgment of the wood engraver is different. His rendering of the composite has made it exactly like one of its components, which it must be borne in mind he had never seen. It is just as though an artist drawing a child had used a portrait closely resembling its deceased father, having overlooked an equally strong likeness to its deceased mother

Wistfully, Galton adds

I trust that the beauty of the woodcut will not be much diminished by the necessarily coarse process of newspaper printing.

T.L.

by exaggerating its shape differences from the sample composite.

This study makes the point that averageness is not the only determinant of attractiveness, and it provides more key evidence of cross-cultural agreement on what is attractive. Averageness, then, may be one of several factors that contribute to beauty, just as texture gradients, occlusion and a host of other cues allow us to perceive depth. The results of Perrett and colleagues, however, tell us little about exactly when non-average physical features contribute to attractiveness.

The 'psychophysics' of beauty has been given a second look by Symons¹². He has reviewed the medical literature on the physical signs of disease and fertility as well as a variety of observations on signs of beauty in various cultures. He sums up the correlations by suggesting that many attractive physical features of female faces (such as a relatively short lower face, gracile jaw, full lips, lighter-than-average unblemished skin, high cheekbones) may be external cues that a female is healthy and fertile, and capable of bearing healthy children. Symons is not suggesting that any man is actually conscious of the evolutionary rationale behind his aesthetic reactions, just that these are the pressures that shaped those reactions as the human brain evolved. Symons has not yet examined the psychophysics of females' perception of the attractiveness of males, but he suggests that some of the cues will probably be the same (for instance, external signs of disease-resistance), others different (for instance, signs of youth).

Clearly, the face of beauty, something we can recognize in an instant, is still difficult to put into words. Although evolutionary psychology has not yet been able to determine the exact face of beauty, it does seem that, as Symons puts it, beauty may be in the adaptations of the beholder. □

Nancy L. Etcoff is in the Neuropsychology Laboratory, Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts 02114, USA.

Safer solutions for chemists

K. P. Johnston

HEAT a compressed gas beyond its critical point, and the gas and liquid phases will become indistinguishable, merging into one 'supercritical' fluid (SCF) phase. Reactions in supercritical fluids have been practised for decades — the polymerization of ethylene, the reaction of ethylene and triethylaluminium to produce olefins, and hydrothermal reactions to form inorganic materials are good examples. But a new generation of reactions in such fluids has been spawned by advances in the fundamental understanding of the nature of SCFs over the past decade¹. The fluids now coming into their own, carbon dioxide (critical temperature $T_c = 31^\circ\text{C}$) and water ($T_c = 374^\circ\text{C}$), are much more environmentally friendly than typical organic solvents. On page 231 of this issue, Jessop *et al.* show that formic acid can be produced much more rapidly from CO_2 in the presence of a ruthenium(II)–phosphine complex under supercritical conditions than in conventional organic solvents².

Provided one works in a region where the gas remains highly compressible, the density and hence the dielectric constant of an SCF may be adjusted over a wide range simply by varying the temperature or pressure. This variation in density influences the chemical potential of solutes and thus reaction rate and equilibrium constants. By tuning the density of a single SCF solvent, more direct information can be obtained about a reacting system than by the conventional technique of studying the reaction in a variety of chemically different solvents. Not only is this tunability valuable for understanding solvent effects on chemical phenomena, but it may be used practically to control phase behaviour, separation processes such as SCF extraction, and rates and selectivities of chemical reactions.

SCFs have a larger free volume and are much more compressible than liquids. Consequently, nonpolar and especially polar attractive forces cause local concentrations of solvent molecules about a given solute to exceed the bulk concentrations by a greater amount than in conventional liquid solution at the same temperature. The augmented local concentrations or clustering can increase hydrogen bonding and ion solvation, and even influence bimolecular rate constants and reaction selectivities³. This unusual behaviour is much more prevalent for reactions with a significant barrier at the transition state than for diffusion-controlled reactions with small barriers.

SCF carbon dioxide has properties very different from conventional hydrophilic and hydrophobic phases. The van der

Waals forces are weak compared with liquid alkanes such as hexane, and resemble those in fluorocarbons. Indeed, DeSimone and co-workers⁴ have synthesized high-molecular-weight perfluoroacrylate polymers in supercritical CO_2 with the same properties as those produced in chlorofluorocarbons. This notable achievement suggests that new CO_2 -based chemical processing could save much of the waste associated with processes that use fluorocarbons or other organic solvents.

New concepts in heterogeneous reactions are being discovered. Surfactants with carbon-dioxide-soluble tails are being developed to form microemulsions and emulsions with either hydrophilic or hydrophobic cores⁵. In another type of heterogeneous reaction, phase-transfer catalysts have now been shown to be capable of transporting ions into an SCF phase (D. Suleiman, D. L. Boatright, C. L. Liotta and C. A. Eckert, manuscript in preparation).

Water undergoes profound changes on heating to the critical point. It expands by a factor of 3, destroying about two-thirds of the hydrogen bonds, and the dielectric constant drops from 80 to 5 (ref. 6). Although a thermodynamic understanding of solutions in supercritical water is emerging, little is known at the molecular level. Recent *in situ* spectroscopic studies with organic probes indicate that pressure and temperature may be used to tune the properties of sub- and supercritical water to mimic that of most organic solvents⁷.

Furthermore, the spectra indicate that clustering of water molecules about ions and polar solutes at supercritical temperatures preserves hydrogen bonding more than might be expected from the bulk density. The hydrophobic effect which limits the solubilities of certain organics at ambient temperatures disappears as water becomes less hydrogen-bonded and more disordered. Supercritical water acts like a 'non-aqueous' solvent, and it readily dissolves many organics and

- Langlois, J. H. *et al.* *Dev. Psychol.* **23**, 363–369 (1987).
- Perrett, D. I., May, K. A. & Yoshikawa, S. *Nature* **368**, 239–242 (1994).
- Wolf, N. *The Beauty Myth* (Morrow, New York, 1991).
- Barkow, J. H., Cosmides, L. & Tooby, J. (eds) *The Adapted Mind* (Oxford Univ. Press, 1990).
- Pinker, S. *The Language Instinct* (Morrow, New York/Allen Lane, London, 1994).
- Ekman, P., Friesen, W. V. & Ellsworth, P. *Emotion in the Human Face* (Cambridge Univ. Press, 1982).
- Symons, D. *The Evolution of Human Sexuality* (Oxford Univ. Press, 1979).
- Barash, D. P. *Sociobiology and Behavior* (Elsevier North-Holland, New York, 1982).
- Langlois, J. H. & Roggman, L. A. *Psychol. Sci.* **1**, 115–121 (1990).
- Galton, F. *J. Nature* **18**, 97–100 (1878).
- Cronin, H. *The Ant and the Peacock* (Cambridge Univ. Press, 1991).
- Symons, D. in *Sexual Nature/Sexual Culture* (eds Abramson, P. R. & Pinkerton, S. D.) (Univ. Chicago Press, in the press).

- Kiran, E. & Brennecke, J. F. *Supercritical Fluid Engineering Science: Fundamentals and Applications* (Am. Chem. Soc. Symp. Series 514, Washington DC, 1993).
- Jessop, P. G., Ikariya, T. & Noyori, R. *Nature* **368**, 231–233 (1994).
- Combes, J. R. & Johnston, K. P. *The Influence of Solvent–Solute and Solute–Solute Clustering on Chemical Reactions in Supercritical Fluids*, 31 (Am. Chem. Soc. Symp. Series 488, Washington DC, 1992).
- DeSimone, J. M., Guan, Z. & Elsbernd, C. S. *Science* **257**, 945 (1992).
- Hoefling, T. A., Enick, R. M. & Beckman, E. J. *J. phys. Chem.* **95**, 7127 (1991).
- Shaw, R. W., Brill, T. B., Clifford, A. A., Eckert, C. A. & Franck, E. U. *Chem. Engng News* **69**, 26 (1991).
- Bennett, G. E. & Johnston, K. P. *J. phys. Chem.* **98**, 441 (1994).
- Jobling, M., Howdle, S. M., Healy, M. A. & Poliakoff, M. *J. chem. Soc. chem. Commun.* **18**, 1287 (1990).

even gases such as oxygen.

Because organics, water and oxygen can be mixed in a single phase, waste water containing several per cent of organics may be oxidized rapidly in this fluid with higher energy efficiency and much less air pollution than in conventional incineration⁶. In addition, water might come to replace certain organic solvents for chemical synthesis at temperatures above 200–350 °C, where most organics become readily soluble.

The miscibility of SCFs and gases already forms the basis for many other reactions. SCFs have been applied to a range of organometallic reactions for the synthesis of dihydrogen and dinitrogen complexes, and for C–H activation, both in solution and on surfaces of insoluble polymers⁸. Jessop and colleagues'

demonstration² of the high performance of formic acid production from CO₂ and H₂ is the latest example, useful as a way of regenerating organic compounds and storing hydrogen. Should any further recommendation be needed, note that the diffusion coefficients are several orders of magnitude higher in supercritical fluids than in liquids.

Given all the advantages in phase behaviour, the ability to tune density to control the outcome, and the benefits of waste minimization with environmentally acceptable solvents, scientists have every incentive to find further uses for these remarkable fluids. □

K. P. Johnston is in the Department of Chemical Engineering, University of Texas, Austin, Texas 78712, USA.

DEVELOPMENTAL BIOLOGY

Unanimity waits in the wings

Patrick H. O'Farrell

EMBRYOLOGICAL development of intricate biological pattern is a multistep process. Beginning with initial asymmetries in the egg, each step incrementally adds detail to a molecular code of positional information — a code that directs the development of structures in their appropriate places. Hedgehog is a signalling molecule that acts at several steps in this process, and two new papers — one in *Cell*¹ and one on page 208 of this issue² — tell us a great deal about this protein's activity in *Drosophila*. Not only does the work reveal remarkable parallels with the action of Hedgehog homologues in vertebrates^{3–5}, but it also points to a link that would unify two schools of thought, each emphasizing different views of the code of positional information.

According to one school, supported by molecular genetic studies in *Drosophila*, positional information is encoded by developmental regulators expressed in localized patterns. Like a paint-by-number kit, where contours delimit areas whose colour is indicated by a number code, domains of gene expression delimit territories of the embryo in which cell fate is specified by a code of regulatory molecules.

In contrast to this mosaic of regulators, experimental embryology suggests that pattern in a variety of organisms is encoded by graded signals. Organizing centres, which are identifiable as transplanted bits of tissue that direct surrounding cells to adopt particular fates according to their distance from the transplant, are thought to emit signalling molecules called morphogens. The idea here is that the fates of cells are dictated by morphogen concentration, which is graded

with distance from the organizer.

The phenotype of *Drosophila* mutant in the *engrailed* gene, which is expressed in segmentally repeated territories in the early embryo, stimulated the original proposal that pattern is specified by a mosaic of territories expressing different developmental regulators^{6,7}. Heemskerk and DiNardo¹ now show that territories of *engrailed* expression act as organizers to specify the fates of surrounding cells, and that Hedgehog, which is produced by *engrailed*-expressing cells^{8–10}, is the morphogen mediating this patterning activity. Hedgehog acts at least twice during the cascade of interactions producing the segmental subdivisions. Mutations in *hedgehog* reveal that Hedgehog signalling is required early for the maintenance of Wingless expression in cells anterior to the *engrailed* territory. To identify the second role, Heemskerk and DiNardo had to distinguish the direct effects of Hedgehog from those due to the lack of Wingless, an important patterning molecule in its own right.

In an elegant piece of work, they show that the production of several distinct types of epidermal cells is independent of Wingless, but dependent on Hedgehog. Using a temperature-sensitive allele of *hedgehog*, they show that the time at which Hedgehog activity is required to specify segmental cell fate (6–9 hours) is distinct from the time at which Hedgehog signals to *wingless*-expressing cells (3–6 hours). Importantly, by genetically manipulating levels of Hedgehog, they provide convincing evidence that Hedgehog acts as a morphogen whose concentration specifies which cell type will form at different positions posterior to the *en-*

grailed stripe. Thus, Hedgehog signalling by the *engrailed* territory suggests a link between the gradient concept of positional information and the code of localized gene expression.

This relation between Hedgehog production and the territory of *engrailed* expression is maintained when the limbs are patterned during the larval-to-pupal transition. As in the embryo, Hedgehog is produced by *engrailed*-expressing cells, which comprise the posterior compartment of the wing and leg disk primordia. In an impressive series of experiments, Basler and Struhl² dissected the role of Hedgehog in limb patterning. They exploited an ingenious technique in which the *flp* recombinase of yeast is used to induce clones that express Hedgehog constitutively (see Fig. 1 on page 208). When these clones occur in the posterior compartment, where the cells naturally produce Hedgehog anyway, they have no effect on the subsequent morphological pattern. But when they occur in the anterior compartment, *hedgehog*-expressing clones can promote dramatic pattern duplications. The most extreme of them — a new wing blade composed of two anterior compartments symmetrically displayed around the clone of *hedgehog*-expressing cells — bears striking similarities to the pattern duplications seen in chick wing upon ectopic expression of the vertebrate homologue, *Sonic hedgehog*⁵.

The pattern duplications demonstrate that Hedgehog, which ordinarily emanates from the posterior compartment, can organize the pattern of the anterior wing blade. However, this patterning action of Hedgehog appears to be indirect. Another signalling molecule (Dpp), the product of the *decapentaplegic* gene (*dpp*), is required for wing formation. Dpp is ordinarily expressed in a stripe of cells just anterior to the posterior compartment. Basler and Struhl show that Hedgehog is required for this expression of Dpp, and that clones that ectopically produce Hedgehog induce Dpp. Indeed, they suggest that the only important function of Hedgehog is to turn on Dpp, an inference which is based on the finding that loss of *dpp* results in the same wing phenotype as loss of *hedgehog*. Thus, Dpp is downstream of Hedgehog in a signalling cascade.

The situation in the leg is similar, but somewhat more complex. As in the wing, ectopic Hedgehog can induce duplications, but the limits of many of these duplications suggested that the dorsal half of the leg differs from the ventral half. Indeed, although Hedgehog induces Dpp in a stripe, this stripe is intense dorsally and faint in the ventral half of the leg disk, where Hedgehog induces another signalling molecule, Wingless. Like Dpp, Wingless is essential for limb patterning. Furthermore, clones that ectopically

express *wingless* support a proposal that Wingless is the intermediary of Hedgehog action in the ventral part of the leg¹¹. Presumably, responses of dorsal and ventral cells to Hedgehog differ because other genes distinguish these two territories.

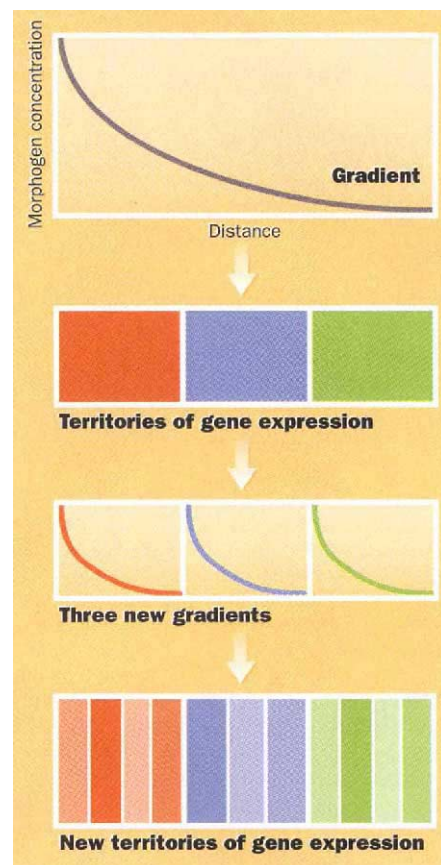
As emphasized, both embryonic and larval Hedgehog signalling emanate from an *engrailed*-expressing territory. What is the importance of this linkage between a territory of localized gene expression and production of a morphogenetic signal? Neither of the original concepts — patterning by a mosaic of localized regulators, and patterning by the action of graded morphogens — fully describes the progressive refinement of pattern that characterizes embryogenesis. By itself, the mosaic of regulators can represent pattern but it is static; and although interpretations of a graded signal can give a step in pattern formation, the model cannot mimic the progressive nature of the process unless one adds a mechanism to specify new organizers. But combining the models overcomes the limitations of either alone (see figure), and the validity of such a combination is supported by the finding that *engrailed* territories of gene expression act as the source of a morphogen.

The power of a combination of mosaic and graded codes has been documented for the patterning of the early syncytial *Drosophila* embryo. Imprints of maternal pattern present in the egg promote accumulation of regulators (Bicoid, for example) in concentration gradients that specify the embryonic axes. These long-range gradients provide coarse positional information that guides expression of the gap gene RNAs in sharply defined, striped territories. The proteins produced by these RNAs diffuse to create local gradients, each steeper and more localized than the primary gradients that define the body axes. This second round of gradients provides local information that guides new gene expression in more finely divided stripes, the pair-rule stripes (territories). And so the process continues, each step refining pattern in the developing embryo¹². In each step, territories of localized gene expression serve as organizers — sources of morphogens — and the gradients of morphogens define new territories. I refer to this alternation between mosaic and graded codes of pattern as a sequential subdivision model; all such models require communication between territories, but in some formulations signalling does not use graded morphogens^{13,14}.

Clearly, the molecular mechanisms involved in communication between different parts of the embryo must change when the embryo makes the transition from syncytial to cellular development. For this reason, there has been little attempt to generalize the lessons learned

from early stages of *Drosophila* development. However, the new results^{1,2} suggest that sequential subdivision continues into later stages.

In Heemskerk and DiNardo's work¹ we see that a territory of gene expression created by the early syncytial cascade now



The sequential subdivision view of pattern formation brings together two modes of encoding positional information, gradients of morphogenic molecules and mosaics of localized gene expression, in an alternating sequence that creates new pattern. As illustrated, a primary gradient of a morphogen across an egg can specify three different territories of gene expression (red, blue and green) as a result of concentration-dependent effects. These three territories can direct formation of new gradients, each of which can direct formation of new territories of gene expression. This sequence can be repeated to specify pattern with increasing accuracy.

acts as an organizer secreting a Hedgehog morphogen. Furthermore, Basler and Struhl's observations² resemble the steps of sequential subdivision. A territory of localized gene expression, the posterior compartment, serves as the organizer producing a signalling molecule, Hedgehog. In the wing, Hedgehog induces a new territory composed of cells transcribing *dpp*. This new territory is the organizer for another patterning signal, Dpp itself. But this may not be the end of the story. Because it is unlikely that the extraordinary detail of final wing pattern is directly specified by graded Dpp concentrations.

further rounds of sequential subdivision may well precede specification of individual cell fates.

Finally, the beautiful analyses of the expression of *Sonic hedgehog* in the chick wing-bud⁵ invite comparisons with the results in *Drosophila*. In fly and chick, Hedgehog signals are produced in the posterior of the developing wing field, and in both ectopic expression induces duplications of more anterior patterns. It seems likely that *Sonic hedgehog* is itself a step in a cascade of sequential subdivision. There are proposals that the limb organizer (the zone of polarizing activity), and perhaps *Sonic hedgehog* expression, depend on expression of FGF-4 by the overlying apical ectodermal ridge¹⁵. The zone of polarizing activity and *Sonic hedgehog* are thought to induce Hox gene expression, but it is not clear whether the Hox genes respond directly to *Sonic hedgehog* concentrations⁵. More likely there are intermediaries in this action and, by analogy with Dpp and Wingless in *Drosophila*, other members of the TGF- β and Wnt families are attractive candidates for intermediate signals. Furthermore the different territories of Hox gene expression are not likely to be the end of the patterning cascade: these domains are too coarse to be responsible for the complete patterning of a limb.

As well as its role in patterning of the limb, *Sonic hedgehog* is expressed in several different regions that have organizer activity, including the notochord, the floorplate of the neural tube and the zone of polarizing activity³⁻⁵. It is possible that *Sonic hedgehog* activates different downstream cascades in each of these cases. But whatever the molecular details, the insights gained from the *Drosophila* studies are likely to throw new light on what may prove to be a general mechanism of embryonic patterning. □

Patrick H. O'Farrell is in the Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA.

1. Heemskerk, J. & DiNardo, S. *Cell* **76**, 449–460 (1994).
2. Basler, K. & Struhl, G. *Nature* **368**, 208–214 (1994).
3. Echelard, Y. *et al.* *Cell* **76**, 1417–1430 (1993).
4. Krauss, S., Cincorden, J.-P. & Ingham, P.W. *Cell* **76**, 1431–1444 (1993).
5. Riddle, R.D., Johnson, R.L., Laufer, E. & Tabin, C. *Cell* **76**, 1401–1416 (1993).
6. Garcia-Bellido, A. *Genetic Control of Wing Disc Development in Drosophila* (Associated Scientific Publishers, New York, 1975).
7. Morata, G. & Lawrence, P.A. *Nature* **255**, 614–617 (1975).
8. Lee, J.J. *et al.* *Cell* **71**, 33–50 (1992).
9. Mohler, J. & Vani, K. *Development* **115**, 957–971 (1992).
10. Tabata, T., Eaton, S. & Kornberg, T.B. *Genes Dev.* **6**, 2635–2645 (1992).
11. Struhl, G. & Basler, K. *Cell* **72**, 526–540 (1993).
12. Lawrence, P.A. *The Making of a Fly* (Blackwell Scientific, Oxford, 1992).
13. Scott, M.P. & O'Farrell, P.H. *A. Rev. Cell Biol.* **2**, 49–80 (1986).
14. Martinez-Arias, A. *Trends Genet.* **5**, 262–267 (1989).
15. Niswander, L. *et al.* *Cell* **75**, 579–587 (1993).

Getting off to a rocky start

Steven V. W. Beckwith

ELEMENTS heavier than helium make up no more than 2 per cent of all baryons, which themselves account for perhaps only 10 per cent of the Universe. Yet they are the building blocks of life, and the separation and build-up of tiny amounts of these heavy atoms into the planet Earth created the laboratory in which complex biochemistry could first work its magic. In an article published this week, Mannings and Emerson boldly present evidence that this build-up is taking place around four other young stars within 500 light years of the Sun¹. If their interpretation is sustained, it will come as a welcome relief to astronomers and planetary scientists who have investigated the general theory of planetary birth for almost 200 years with no examples to study directly.

The evidence is imprinted in the radiation from small solid particles orbiting the stars. The dust contains raw materials for planet-building in the form of carbon, silicon, iron and other 'metals' (in astronomers' parlance, any element heavier than helium). Surrounding each star, the orbits lie preferentially in a thin plane called a 'disk', keeping the particles close together where they collide and occasionally stick together, growing from submicrometre-sized dust grains to small pebbles. Eventually, the pebbles collect into much larger planetesimals, whereupon gravity assists the building process to assemble planets.

The growth changes their radiation properties: micrometre-sized particles

radiate effectively in the optical band but very poorly in millimetre waves, where the wavelength is much larger than the radiating surfaces. Small pebbles a few

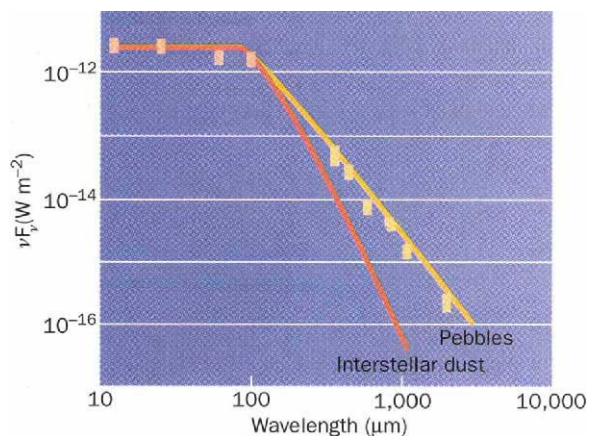


FIG. 1 Spectral energy distribution of the star DG Tau alongside theoretical curves for radiation by interstellar dust grains (red) and millimetre-sized particles (orange). The yellow points are new measurements by Mannings and Emerson¹; the pink points are from the IRAS database, where the disks are opaque.

millimetres across produce relatively more of the long-wavelength light, the change being a factor of 10 or more. Such enhancements are easily detectable, in principle. Mannings and Emerson searched for this change by observing the spectra of six disks between 350 μm and 2 mm and comparing their radiative properties with calculations made for rocky material during a planetary build-up². Four of the disks give acceptable matches to the calculated spectral energy distributions; all four are unquestionably at odds with the radiative properties of normal interstellar dust grains (see Fig. 1).

It has been only 10 years since pre-planetary disks were first observed³ and less than 5 years since large samples were identified for serious study⁴. In the interim, the masses, temperatures and sizes of the disks have been measured and compared with the minimum requirements to assemble planetary systems like our own. The disks are surprisingly common, accompanying the birth of nearly half of the nearby solar-mass stars. They are also cool; much of their radiation emerges at far-infrared and submillimetre wavelengths where the Earth's atmosphere is opaque. In fact, there was

evidence for such disks almost 20 years ago, when infrared detectors first became sensitive enough to detect the disk radiation as an excess over the stellar photosphere in nearby T Tauri stars, the youngest visible counterparts to the Sun⁵. But alternative explanations were more plausible at the time, and the disk interpretation did not gain favour until far-infrared measurements became available. Then, in 1983, the IRAS satellite produced the first far-infrared sky survey showing clearly that the radiation from T Tauri stars arose in disks of small particles. Remnant disks are seen around many main-sequence stars, too, the most famous being Vega, the brightest star in the constellation Lyra.

Mannings and Emerson's work is the latest in a series of attempts to find evidence for particle growth via long-wavelength spectral signatures⁶. The main advantage of the new study is its very broad spectral coverage — a factor of 6 in wavelength — and its analysis of the resulting spectra taking into account the opacity of the disks. To measure the radiative properties of particles in the disks reliably, it is important that one sees *all* the particles. If the total opacity is too high, as is usually the case at wavelengths shorter than 100 μm , one observes only thermal black-body radiation independent of the properties of the particles themselves. Almost all the disks around the very young stars (in which particles should still be growing) have regions of high opacity. The trick is to select those in which the disk opacity is small and remove residual effects with a careful modelling of the disk structure. The accuracy of the modelling increases when the data set contains many measurements over a large wavelength range.

Unfortunately, particle growth is not the only way to produce the spectra seen from these stars. Changes in particle shape can also enhance the millimetre-wavelength radiation. An especially intriguing idea is that particle growth occurs, but the growth is initially of a fractal character⁷. The resulting particles resemble snowflakes or the long, thin tendrils of a leafy plant. The tendrils act as antennae, coupling well to the long-wavelength radiation without using much additional material. Fractal particles may even have been necessary in the primitive solar nebula to make the solids more responsive to the gas and allow efficient growth⁸. The interplanetary particles captured by high-flying aircraft⁹ more closely resemble fractals than the idealized hard spheres adopted in theories of interstellar grains (see Fig. 2).

Chemical reactions in the dense pre-planetary soup could also have enhanced the grain radiation to mimic particle growth. But it is difficult to see how compositional changes in the grain mat-

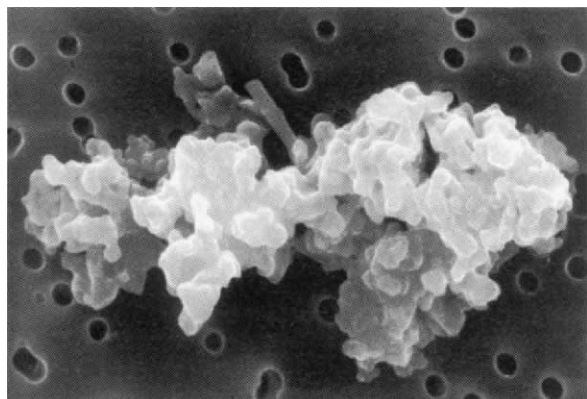


FIG. 2 Micrograph of an interplanetary particle collected from the Earth's stratosphere, showing the departure from spherical symmetry⁹. Particle is about 10 μm across. (Courtesy of D. E. Brownlee.)

erial can create the strong effects seen in the spectra of these disks. At 2 mm wavelength, some of the disks emit 30 times the radiation expected from normal interstellar dust particles. Radio radiation is strengthened far more effectively by changing the size and shape of the antenna than by changing its chemical makeup.

Whatever is responsible for the enhanced radiation, there is no doubt that the solid-state constituents in these disks are markedly different from the interstellar dust whence they came. Whether the result of growth or chemical processing, there is strong evidence that the solids are being transformed as they orbit the stars. Such transformations are precisely what we need to see to determine whether the particles make their way into planets. Even if the present work needs more

support to define unambiguously the onset of planet formation, we now have a sample of stars for testing our theories, and testable criteria by which we can explore the processes that shaped our own rocky beginning from a dusty past. □

Steven V. W. Beckwith is at the Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany.

1. Mannings, V. & Emerson, J. P. *Mon. Not. R. astr. Soc.* **267**, 361–378 (1994).
2. Miyake, K. & Nakagawa, Y. *Icarus* (in the press).
3. Aumann, H. H. et al. *Astrophys. J.* **278**, L23–27 (1984).
4. Strom, K. M. et al. *Astr. J.* **97**, 1451–1470 (1989).
5. Lynden-Bell, D. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **168**, 603–637 (1974).
6. Beckwith, S. V. W. & Sargent, A. I. *Astrophys. J.* **381**, 250 (1991).
7. Wright, E. L. *Astrophys. J.* **320**, 818–824 (1987).
8. Meakin, P. & Donn, B. *Astrophys. J.* **329**, L39–42 (1988).
9. Brownlee, D. E. et al. *Proc. 8th Lunar Sci. Conf.*, 149 (Pergamon, New York, 1977).

ECOLOGY

Flummoxed by fake fruits

Peter D. Moore

MANY plants depend on animals for the dispersal of their fruits and seeds, and such species usually advertise their wares by the use of bright colours and conspicuous forms. But for others the levels of seed predation can become unacceptably high and selective pressures operate in the reverse direction, forcing the plant to hide or disguise its fruit. One of the most unusual methods of achieving this end is described by P. K. Groom, B. B. Lamont and H. C. Duff in the latest issue of *Functional Ecology* (8, 110–117; 1994) — they have identified a plant in which specialized leaves that resemble fruits serve to confuse the birds that seek to strip the plant of its produce.

Plant leaves are remarkably adaptable. Besides their obvious photosynthetic role, they may be modified to protect buds, form spines that deter grazers, construct traps that catch insects, make tendrils that enable a stem to climb, or house secretory organs that attract ants. It has even been argued that variegated leaves in patchy shade operate as camouflage, reducing the conspicuousness of the plant in the eyes of potential grazers, and the same may be true of highly dissected leaves and those with holes in their laminae, such as the Swiss cheese plant, *Monstera deliciosa*. But in the Australian shrub *Hakea trifurcata* (Proteaceae) the mature shoots develop distinctive leaves with a form that bears a strong resemblance to the follicles that contain the seeds, and Groom, Lamont and Duff propose that these leaves serve to distract and confuse the main predator of the follicles, the white-tailed black cockatoo.

Hakea trifurcata is unusual among the Australian Proteaceae in that its leaves

are dimorphic. The juvenile plant produces only narrow, needle-like leaves but, on maturity, a second type of leaf develops alongside the needle leaves on the same branches. These are broader, rather

sense in this species because, unlike most members of the Proteaceae, it is unable to resprout after fires and is entirely dependent on seed production for population recovery.

To test the hypothesis that the broad leaves make it more difficult for cockatoos to locate and harvest *Hakea* fruits, caged birds were subjected to feeding trials in which they were offered normal fruiting branches with broad leaves present, and treated fruiting branches with all broad leaves removed. The total number of follicles subsequently removed by the birds from treated branches was significantly higher than in those with broad leaves present. Overall, the birds approached the broad leaves in the untreated branches two to three times more often than real fruits, perhaps being further distracted by the generally larger size of the deceptive leaves, especially those at greater distances from the genuine article. Having developed a search image appropriate for the detection of *Hakea* follicles, the birds are evidently easily led astray by the cryptic leaves.

Mimicry is, of course, quite common in nature, though self-mimicry, or 'self-crypsis' as the authors term it, is most unusual, if not unique, among plants. The evolution of this strategy in *Hakea trifur-*



Spot the fruit — this branch of *Hakea trifurcata* bears two fruits, the rest of the structures being leaves. But which is which? The photograph is about two-thirds life size.

spatulate structures, which in shape and colour (yellow-green) are remarkably like the plant's follicles. The broad leaves are generally somewhat longer than the follicles, but those that are situated closer to the fruits are closer to them in size. They are also denser in those areas where the follicles are present. Protection of the fruits from excessive predation makes

cata demonstrates that the energy invested in deception evidently pays off, and the moral of the story seems to be that if you really want to confuse a predator, then go around looking like yourself. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

B. B. Lamont & P. K. Groom

Cosmic-ray hotspots

A. G. W. Cameron

THE discovery¹ of gamma-rays from excited states of ^{12}C and ^{16}O in the direction of Orion startled many astronomers. Some of the implications are explored by Donald Clayton on page 222 of this issue². In particular, the observations could account for the surprisingly large amount of ^{26}Al in interstellar space, and for other extinct radioactive elements whose daughter products are found in meteorites and which are inferred to have existed in the primitive solar nebula.

The excited states from which the γ -rays are inferred to be emitted are at 4.43 MeV in ^{12}C and at 6.13 MeV in ^{16}O ; the γ -ray source is in the Orion star-forming molecular cloud complex, which is the nearest large star-forming region in our Galaxy. The team who made the discovery¹, using the Compton Telescope on board the Compton Gamma Ray Observatory, concluded that the carbon and oxygen in question were in the form of energetic particles, probably with energies of 3 to 10 MeV per nucleon. The excited states would be populated by nuclear reactions taking place on C, N, O and Ne energetic nuclei when they collided with hydrogen and helium in the molecular clouds of Orion. But the feature that has aroused the interest of many researchers in astrophysics and the planetary sciences is that the fluxes of these energetic nuclei were deduced to be some 30 times greater than would be expected from measurements of cosmic rays near the Sun.

Acceleration

Should this high cosmic-ray flux in Orion be a great puzzle? Probably not, in retrospect. Large star-forming regions make their share of massive stars, usually in groups that become OB associations, which are local clusters of stars observed to expand away from a common point within a cloud. The most massive of these stars evolve quickly (within a few million years, short compared to the lifetime of a molecular cloud complex), and then they explode as supernovae of type II at nearly the same time.

The gas ejected from these supernovae, if there are enough of them, combines with the energetic particles accelerated in their explosions, to form an expanding cloud of hot plasma known as an interstellar superbubble (these superbubbles should make promising targets for the Compton Telescope). On a somewhat longer timescale the less massive stars, down to about 8 solar masses³, continue to explode as supernovae.

So star-forming regions will possess continuing sources of cosmic-ray accelera-

tion, as the time between formations of OB associations should be short compared with the evolution time of an 8-solar-mass star (about 30 million years). The accelerated cosmic rays will diffuse out of the star-forming region, of course, but that too will take time, and thus star-forming regions should have higher fluxes of cosmic rays than the ordinary interstellar medium.

Probably some brave theorists will attempt to estimate these diffusion times, but the magnetic fields in star-forming regions must form a complex tangle, and this calculation will not be easy to do. More important, the supernova accelerators are themselves some of the main producers of heavier elements, so it should not be surprising if the locally produced cosmic rays are enriched in elements such as C, O, Ne and beyond. The Compton Observatory has opened up a fascinating new field of research.

As it is not currently feasible to estimate how the cosmic-ray flux varies with time in star-forming regions, Clayton has taken the simple approach of assuming that the Orion's cosmic-ray flux should also exist in other star-forming regions. He has further assumed that the fluxes of the more massive cosmic rays are enhanced by the same factor as C, N, O and Ne. These simple assumptions lead to far-reaching conclusions and are a good way to begin, but may well lead to errors of a factor of a few. Orion does not currently contain an obvious interstellar superbubble, so Clayton may have underestimated the real time averages of the effective fluxes in star-forming regions.

Clayton's first conclusion concerns the production of ^{26}Al in the interstellar gas. There is a total of about 2 solar masses of this radioactive nuclide (half-life about 0.7 million years) in the galactic gas, with a concentration towards the Galactic Centre. This has long been a puzzle, as the amount has seemed considerably to exceed what could reasonably be produced by novae, supernovae and stars in the appropriate phase of evolution at their current rates of occurrence in the Galaxy. But by Clayton's assumptions the interstellar ^{26}Al can easily be produced by cosmic rays in star-forming regions.

His second conclusion deals with the extinct radioactive nuclides whose daughter products reside in certain mineral phases in primitive unequilibrated meteorites, and which are believed to have been live at the time of formation of meteoritic parent bodies in the primitive solar nebula.

Previously it has seemed necessary to attribute these radioactivities to nucleo-

synthesis taking place in stars close to the site of formation of the primitive solar nebula, so that in principle the nuclides could be transported from their sources to the primitive solar nebula in a few of their half-lives. In fact, it was thought that the shorter-lived radioactive nuclides could mostly be attributed to nucleosynthesis in an asymptotic giant branch (AGB) star late in its evolution^{4,5}. The AGB star, so the idea goes, would have triggered the formation of the primitive solar nebula in a star-forming region. Some problems have developed with this picture, not least of which is that AGB stars seem scarce in star-forming regions⁶, so the formation of the primitive solar nebula must have been a near-unique event. This prospect has generated a certain amount of unease.

Abundances

But Clayton has estimated the rate of production of the shorter-lived (and some of the longer-lived) extinct radioactive nuclides by cosmic rays in star-forming regions. There are many uncertainties, of course, but it seems plausible that the shorter-lived radioactivities in the primitive solar nebula could have been produced with the required abundances in this way.

In fact, the cosmic rays would continue to produce these radionuclides until the final stages of the collapse of a core (a dense site of star formation) in the cloud. So all stars could have such extinct radioactivities present in the circumstellar accretion disks from which planets may have been made, and the Solar System would not be unique. If cosmic-ray bombardment of the gas can continue during most of the period of formation of the solar nebula, then even very short-lived radionuclides might produce detectable isotopic anomalies in meteorites.

Actually, Yin and co-workers believe that they have detected one such short-lived extinct radioactivity in meteoritic materials, ^{99}Tc , with a half-life of only 0.2 million years⁷. If confirmed, their observation would pose an acute problem for the AGB star theory, but could readily be accommodated by the new theory of production by cosmic rays. This should be a spur, if any were needed, to new measurements. □

A. G. W. Cameron is in the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

1. Bloemen, H. et al. *Astr. Astrophys.* (in the press).
2. Clayton, D. D. *Nature* **368**, 222–224 (1994).
3. Kennicutt, R. C. Jr *Astrophys. J.* **277**, 361–366 (1984).
4. Cameron, A. G. W. in *Protostars and Planets III* (eds Levy, E. H. & Lunine, J. I.) 47–73 (1993).
5. Wasserburg, G. J., Busso, M., Gallino, R. & Raiteri, C. M. *Astrophys. J.* (in the press).
6. Kastner, J. H. & Myers, P. C. *Astrophys. J.* **421**, 605–614 (1994).
7. Yin, Q., Jagoutz, E. & Wänke, H. *Meteoritics* **27**, 310 (1992).

Tracking neurotrophin function

Alun M. Davies

THE neurotrophins were originally identified by their ability to prevent the death of developing neurons. But study of them has greatly influenced thinking about how cell survival is regulated in a wide variety of tissues, which is why four recent papers¹⁻⁴, including two on pages 246 and 249 of this issue^{3,4}, will attract general attention. The papers describe the use of gene targeting to generate mice with null mutations in genes encoding each of the signal-transducing receptors for neurotrophins and in one of the neurotrophin genes. Each of the mutations results in distinctive and striking defects in anatomy and behaviour. Not least, this work paves the way for definitive analysis of the function and specificity of neurotrophin receptors.

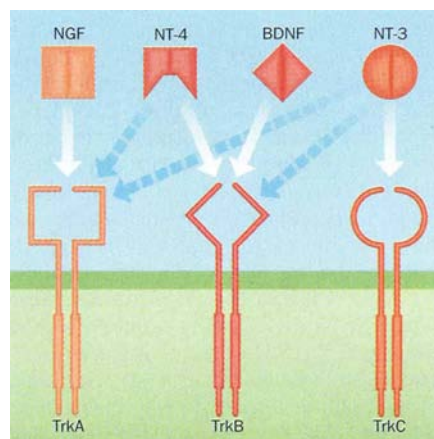
A paradoxical feature of the developing vertebrate nervous system is the occurrence of widespread neuronal death. This takes place in many populations of neurons during a restricted period of development shortly after neurons innervate their targets, and serves to match neuronal numbers to target field requirements. The proposal by Levi-Montalcini and Hamburger that target fields regulate their innervation density by the provision of limiting quantities of neuron survival factors, for which the innervating neurons compete, was substantiated by the identification of the first neurotrophin, nerve growth factor (NGF), four decades ago⁵. The best evidence for this proposal was that developing neurons whose survival is promoted by NGF *in vitro*, namely sympathetic and some sensory neurons, also depend on NGF *in vivo*. Anti-NGF antibodies administered during the phase of target field innervation largely eliminate these neurons whereas injection of NGF *in vivo* rescues neurons that would otherwise die.

A second neurotrophin, brain-derived neurotrophic factor (BDNF), was purified by Barde a decade ago, and its subsequent cloning demonstrated that it is highly related to NGF⁶. This discovery rapidly led to the cloning of neurotrophin-3 (NT-3) and NT-4 (also called NT-5 or NT-4/5). Studies *in vitro* indicate that each neurotrophin promotes the survival of a characteristic set of developing neurons, although some neurons can be supported by more than one neurotrophin⁶.

Neurotrophins bind to two kinds of transmembrane glycoproteins, p75 and members of the Trk (pronounced track) family of receptor tyrosine kinases⁷. p75 binds all neurotrophins with similar affinity but does not seem to be a functional receptor in the absence of a Trk receptor. It is nonetheless important for the de-

veloping nervous system because mice bearing a null mutation in the p75 gene have decreased pain sensitivity and cutaneous innervation⁸, and cultured sensory neurons from these mice are less sensitive to NGF than wild-type neurons⁹. But how p75 specifically enhances the survival response of sensory neurons to NGF is not understood.

The Trk receptor tyrosine kinases bind neurotrophins with higher affinities than



Binding preference of homodimeric neurotrophins to Trk receptor tyrosine kinases. The preferred interactions are illustrated by solid arrows and the non-preferred interactions by dashed arrows.

p75, distinguish between neurotrophins and are essential for functional responses *in vitro*⁷. Three members of the family have been identified in vertebrates: TrkA, TrkB and TrkC. Expression of these receptors in fibroblasts has been useful for investigating neurotrophin-receptor relationships, and it seems that TrkA is the preferred receptor for NGF, TrkB is the preferred receptor for BDNF and NT-4, and TrkC is the receptor for NT-3 (see figure). These studies have also raised the possibility of cross-talk between NT-3 and TrkA and TrkB, and there is controversial evidence that NT-4 may be able to elicit responses via TrkA.

A further complex feature of Trk receptor tyrosine kinases is the existence of several isoforms. In particular, the *trkB* and *trkC* genes code for isoforms that lack the tyrosine kinase domain. Whether these non-catalytic receptors couple to unknown signalling pathways or modulate the response of neurons to neurotrophins by competing for ligand is not known.

Although the administration of neutralizing anti-NGF antibodies to developing vertebrates has been valuable in establishing the developmental role and neuronal specificity of NGF⁵, this approach has its limitations. Hence the

adoption of gene-targeting techniques to study the neurotrophins and their receptors, as described in the four recent papers¹⁻⁴.

The *trkA* mutant mice were generated by deleting sequences encoding the kinase domain³. Most homozygous mutant mice died within a month of birth, although some reached adulthood. These mice were insensitive to temperature and pain, and had selective loss of small-diameter neurons in spinal and trigeminal ganglia that convey these sensations. Paravertebral sympathetic neurons were also virtually absent by the tenth postnatal day. This phenotype is very similar to that of mice exposed to anti-NGF antibodies *in utero*, and is entirely consistent with the neuronal specificity of NGF observed *in vitro*.

The new findings formally demonstrate that the trophic actions of NGF are mediated by TrkA *in vivo*, and they suggest that the action of NT-4 and NT-3 on TrkA observed in cell lines is not relevant for the developing peripheral nervous system. If this cross-talk were physiologically relevant, one would expect anti-NGF antibodies to be less effective in eliminating sympathetic and sensory neurons than the *trkA* null mutation. Detailed comparison of the phenotype of the *trkA* mutants with that of a mouse homozygous for a null mutation in the NGF gene should further clarify matters, and would reveal if TrkA and NGF have additional functions independent of each other.

There is evidence that, as well as sensory and sympathetic neurons, NGF stimulates certain cholinergic neurons in the central nervous system (CNS). NGF regulates the expression of choline acetyltransferase (ChAT) in subsets of cholinergic neurons of the basal forebrain and prevents death of these neurons following axotomy¹⁰. In the *trkA* mutants the level of ChAT is reduced in these neurons but their number is apparently unaffected³. So whereas endogenous NGF regulates ChAT expression in basal forebrain cholinergic neurons, it seems that it is not normally required for survival. Interestingly, ChAT staining of the striatal cholinergic neurons that normally express TrkA is unaffected in the *trkA* mutant. It will be important to ascertain if the absence of a functional TrkA receptor causes some phenotypic abnormality in these neurons; if it does, it would imply that NGF has other functions in the CNS.

The *trkB* mutant mice were generated by deleting sequences encoding the kinase domain, which should leave the non-catalytic isoforms unaffected¹. The mice died within a few days of birth and had reduced numbers of sensory and motor neurons. These findings are consistent with the observation that the two preferred TrkB ligands, BDNF and NT-4, promote the survival of motor neurons *in*

*vitro*¹¹ and subsets of sensory neurons¹².

Comparison of the phenotypes of the *trkB*¹ and BDNF² null mutants provides an indication, albeit indirect, of the relative importance of BDNF and NT-4 in promoting the survival of different neurons *in vivo*. Analysis of the sensory ganglia of postnatal BDNF mutant mice² reveals marked loss of neurons in those ganglia that have been shown *in vitro* to contain a large proportion of BDNF-dependent neurons. Similar numbers of sensory neurons are lost in the dorsal root ganglia of *trkB* and BDNF null mutants (30%) but more neurons are lost in trigeminal ganglia of *trkB* (60%) than BDNF mice (44%). The most striking difference is that lack of TrkB receptors causes marked reduction in motor neuron number (70% in the facial nucleus and 35% in the lumbar cord), whereas mice without BDNF are normal in that respect. These findings suggest that, *in vivo*, BDNF is important for the survival of certain sensory neurons but not motor neurons, and raise the possibility that NT-4 promotes motor neuron survival. The relative roles of BDNF and NT-4 would be clarified by studying the phenotype of mutant mice that lack NT-4. Moreover, analysis of the phenotype of intercrosses between BDNF and NT-4 null mutants should reveal if TrkB has functions *in vivo* that are additional to the combined actions of its two preferred ligands.

An intriguing finding in postnatal BDNF mutants is that most, if not all, of the peripheral axons of the remaining vestibular neurons terminate in connective tissue short of their target epithelia², raising the possibility that BDNF may be involved in axonal guidance. But because the peripheral projection of the neurons was examined postnatally it is not known whether they had failed to reach their targets or whether they had done so and subsequently retracted. This matter of the developmental stage at which observations are made highlights a general need for caution in interpreting the results of these gene-targeting experiments. Because mutations in the neurotrophin and *trk* genes may have developmental consequences very early on, the reduced size of neuronal populations in mutant mice could result not just from the failure of neurons to survive but from aberrations in the proliferation and differentiation of neuronal progenitor cells. This is especially germane given that, *in vitro*, NT-3 can promote proliferation of neural crest cells¹³ and BDNF can commit such cells to the sensory neuron lineage¹⁴.

Finally, the *trkC* null mutants. These mice were generated by deleting sequences encoding the kinase domain, leaving non-catalytic isoforms unaffected⁴, and most died within three weeks of birth. Notably, though, they

displayed abnormal motor behaviour suggestive of defective proprioception (sensory information about limb movement and position). Consistent with that observation, the mice have 20% fewer neurons in dorsal root ganglia and selectively lack group Ia afferents, which innervate muscle spindles. These results are in accord with the demonstration that NT-3 promotes the survival of proprioceptive neurons *in vitro*¹⁵. Whether *trkC* mutant mice have deficiencies in any of the other subsets of sensory neurons that respond to NT-3 *in vitro* remains to be seen.

Intriguingly, although *trkB* and *trkC* messenger RNA transcripts are widely expressed in the CNS, initial studies have not revealed any gross defects in the CNS of either *trkB*¹ or *trkC*⁴ mutants. Although abnormalities may yet be identified, the fact that expression of these two genes overlaps in many regions means that they could be involved in functionally redundant pathways. Intercrosses between individual *trkB* and *trkC* mutants should help here. Another issue that will have to be addressed is the role of the non-catalytic TrkA and TrkB isoforms, because in these experiments their coding sequences were unaltered^{1,4}. Mice with disrupted extracellular sequences in *trkB* and *trkC* genes might be informative in this respect.

What next? There is a wealth of detailed anatomical data yet to be obtained from the mutant mice and from intercrosses between them. Moreover, neurons that lack specific Trk receptor tyrosine kinases can now be obtained for study *in vitro*. The contradictory results obtained from investigation of p75 in physiologically irrelevant cell lines illustrate the importance of looking at the function of neurotrophin receptors in the appropriate cellular and developmental context. Thus, analysis of neurons isolated from embryos homozygous for mutations in *trk* genes should improve our understanding of the cell biology of Trk receptors, as illustrated by equivalent studies of neurons from p75 null mutant mice⁹. □

Alun M. Davies is in the School of Biological and Medical Sciences, St Andrews University, St Andrews, Fife KY16 9TS, Scotland.

1. Klein, R. *et al.* *Cell* **75**, 113–122 (1993).
2. Ernfors, P., Lee, K.-F. & Jaenisch, R. *Nature* **368**, 147–150 (1993).
3. Smeyne, R. J. *et al.* *Nature* **368**, 246–249 (1994).
4. Klein, R. *et al.* *Nature* **368**, 249–251 (1994).
5. Levi-Montalcini, R. *Science* **237**, 1154–1162 (1987).
6. Thoenen, H. *Trends Neurosci.* **14**, 165–170 (1991).
7. Meakin, S. O. & Shooter, E. M. *Trends Neurosci.* **15**, 323–331 (1992).
8. Lee, K. F. *et al.* *Cell* **69**, 737–749 (1992).
9. Davies, A. M., Lee, K. F. & Jaenisch, R. *Neuron* **11**, 565–574 (1993).
10. Dreyfus, C. F. *Trends Pharmacol. Sci.* **10**, 145–149 (1989).
11. Henderson, C. E. *et al.* *Nature* **363**, 266–270 (1993).
12. Davies, A. M. *et al.* *J. Neurosci.* **13**, 4961–4967 (1993).
13. Kaicheim, C., Carmel, C. & Rosenthal, A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1661–1665 (1992).
14. Sieber-Blum, M. *Science* **243**, 1608–1611 (1989).
15. Hohn, A. *et al.* *Nature* **344**, 339–341 (1990).

DAEDALUS

Fair copy

COMPUTERS are still very bad at interpreting handwriting. The 'digital personal assistants' sold for this purpose are little more than expensive jokes. Daedalus points out that handwriting is very much the expression of a personal neural code. Those who learn to write with their feet, or with a pen held in their teeth, continue to produce the same characteristic script. So he plans to intercept the neural handwriting code one step nearer its source. He is devising a machine to read, not the writing itself, but the movements that generate it. He points out that our fingers are operated by tendons connected to muscles in the lower arm. A set of sensors attached to the wrist could read the movements of these tendons. They would be closer to the origin of the handwriting; they would also be dealing with a set of purely scalar values rather than a complex two-dimensional pattern.

DREADCO's technicians are exploring several approaches to this scheme. Their simplest idea is a bracelet carrying a set of piezoelectric movement sensors, each responding to the change of tension in a single tendon as it is operated. Another idea is to implant a tiny magnetic speck in each tendon, and follow their movements by a set of pick-up coils on the bracelet. Less invasively, a magnet on the bracelet could induce a tiny potential across each tendon as it moved; skin-voltage sensors could then read its velocity. This differential signal might be more informative than reading the displacement directly. In each case the outputs from the bracelet sensors would be led by a fine wire to a small neural-net learning computer in the wearer's pocket. It would accumulate examples of his tendon movements until it had learnt his internal handwriting code.

The wearer of DREADCO's 'Wristwriter' will enjoy a new scriptive freedom. He need not look at what he is writing, or keep within the confines of a tiny sensing pad, or avoid overwriting his previous comments. He can scribble in bad light or darkness on any scale and on any surface. For confidential notes, he can use an empty pen or even a wooden stick, leaving no trace except the private record in his Wristwriter. Even if he cannot read his own handwriting, it will still be the outcome of a consistent code of tendon movements. The Wristwriter will read it for him.

The Wristwriter will even be able to interpret other manual activities. It will memorize knitting patterns, and convert spontaneous flute tooting and piano playing into musical notation. Wherever fingers are at work, it will be there to record what they do. David Jones

Origin of insect pollination

SIR — The Gnetales, a bizarre order of gymnosperms thought to be a sister group of angiosperms^{1,2}, comprise three distinct genera, *Ephedra*, *Welwitschia* and *Gnetum*; the first two are xerophytic and pollinated mainly by wind and partly by insects³. *Gnetum* is restricted to tropical rain forests and its pollination mechanism has not been definitely established. We have observed insect visits to a dioecious shrub, *Gnetum gnemon*, in a lowland mixed dipterocarp forest in Sarawak, Malaysia. Pollination droplets were secreted from ovules on female strobili and from sterile ovules on male strobili in the evening and were consumed by nectar-seeking moths of Pyralidae and Geometridae. Sticky pollen was attached on probosces and antennae of these moths. The moth attraction in *Gnetum* would have arisen from unspecialized entomophily, which is thought to be an original pollination system of early Gnetales and early angiosperms⁴.

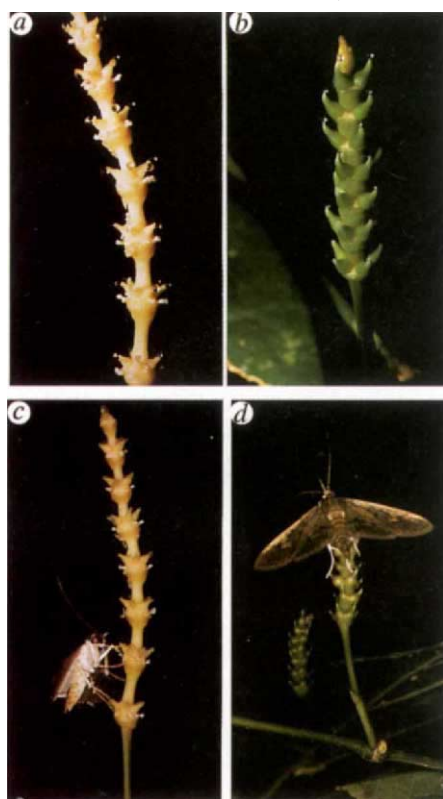
One hypothesis of the angiosperm diversification in the Cretaceous is the adoption of insect pollination⁵. Recent palaeobotanical evidence suggests that early Gnetales and early angiosperms co-occurred in mesic floodplain habitats and had similar vegetative morphology (small herbs or shrubs)⁶, similar ecological tolerances and similar reproductive biology (flowers simple and pollinated by wind or unspecialized nectar-seeking insects)⁴. Thus, pollination systems of extant Gnetales may hold the key to the problem of angiosperm radiation. Among the three genera of the Gnetales, anemophily is prevalent in *Ephedra* and entomophily is known in only a few species of *Ephedra*⁷ and rarely in *Welwitschia*⁸. *Gnetum* is an enigmatic genus: little is known of its pollination biology, although there are some morphological indications that point to entomophily⁹.

In the Malayan region, 16 species of *Gnetum* are known; two are trees or shrubs and others are woody climbers. We studied the floral biology of a dioecious shrub, *Gnetum gnemon* Linné var. *tenerum* Markgraf, in a lowland mixed dipterocarp forest in Lambir Hills National Park, Malaysia (4° 2' N, 113° 50' E, altitude 60 m). We observed flowering throughout the year, but the flowering period of individual shrubs was less than 3 weeks. In preliminary observations made in the daytime between 21 and 30 November 1992, we found no flower visitors. We made night observations between 5 and 31 August 1993.

Ovules on female strobili and sterile ovules on male strobili secreted pollination droplets in the evening. Sequential sampling of droplets of pollinator-excluded strobili showed that droplet

secretion started at about 18:00 h, and ceased by 21:00 h in male and by 20:00 h in female strobili, respectively. If droplets remained intact, the mean standing crop of a droplet increased to 0.164 in male and to 0.190 µl in female strobili at 21:00 h (a, b in the figure). We did not observe rapid withdrawal of droplets after pollination.

The droplets contained sugar, the concentration of which ranged from 3 to 13%, which is lower than that of an entomophilous *Ephedra* (72–80%)⁷ but is higher than that of anemophilous conifers (for exam-



Strobili of *Gnetum gnemon* in the evening in a tropical rain forest in Sarawak. a, Sterile ovules emitting droplets (male strobilus); b, fertile ovules emitting droplets (female strobilus); c, a male strobilus visited by a pyralid moth, *Herpetogramma* sp.; d, a female strobilus visited by a pyralid moth, *Hedylepta* sp.

ple, *Pinus*, 1.2%)⁹. The sugar concentration of the droplet was affected by the relative humidity of the surrounding air, which usually increased to more than 95% in the evening in the forest floor habitat. Nectar was also secreted among microsporangiphores on male strobili, although in very small amounts.

Male strobili secreting droplets had a strong odour, whereas female strobili were less fragrant. Probably attracted by the odour, nocturnal moths of Pyralidae and Geometridae visited the strobili between 18:00 and 21:00 h. We found nine pyralid and four geometrid species. Each moth landed on a strobilus and extended

its proboscis to suck the droplets (c, d in the figure). Moths visiting male strobili sometimes probed the nectar secreted among male flowers and then their probosces touched the microsporangia. Moths fluttered slowly between strobili and the mean time spent on a strobilus was 485 s ($n=5$, range 72–1,155). Moth visits per strobilus per night were more frequent on a male (1.74 ± 0.62 , $n=6$) than a female strobilus (0.67 ± 0.28 , $n=3$). Each male and female strobilus was visited about 8.7 and 3.4 times on average respectively during the flowering period, which is about 5 days.

The pollen grains are inaperturate and sticky while a pollenkitt is absent. All the 18 moths collected on male strobili had pollen grains attached on probosces and antennae and, much less frequently, on tarsi of legs. Although there were no pollen grains attached on a moth collected on a female flower, the same species was collected on male strobili.

Most *Gnetum* species, even tall climbers, flower in the understory of rain forests where the humid habitat is least suitable for anemophily. *Gnetum* strobili, by emitting droplets as a nectar reward and odour as an attractant, attract nocturnal moths. *Gnetum* does not have typical moth flowers in a classical anthecological sense because petals are absent and the nectar reward is exposed. However, the nocturnal pattern of droplet secretion, the low sugar concentration of the droplet and fragrant coincide with the moth-pollinated flower syndrome. High relative humidity and stillness of surrounding air in the evening favour this pollination system, which is also adopted in *Nepenthes* (Nepenthaceae) in the Palaeotropics¹⁰. We believe that the moth attraction in *Gnetum* derived from unspecialized entomophily as observed in *Ephedra* and *Welwitschia* after the divergence of long-tongued lepidopterans which coevolved with deep angiosperm flowers in the late Cretaceous.

Makoto Kato

Biological Laboratory, Yoshida College, Kyoto University, Sakyo, Kyoto 606-01, Japan

Tamiji Inoue

Centre for Ecological Research, Kyoto University, Shimosakamoto, Otsu 520-01, Japan

- Crane, P. *Ann. M. Bot. Gdn* **72**, 716–793 (1985).
- Doyle, J. A. & Donoghue, M. J. *Paleobiology* **19**, 141–167 (1993).
- Kubitzki K. in *The families and genera of vascular plants* vol. 1 (eds Kramer, K. U. & Green, P. S.) 378–391 (Springer, Berlin, 1990).
- Lloyd, D. G. & Wells, M. S. *Pl. Syst. Evol.* **181**, 77–95 (1992).
- Friis, E. M. et al. (eds.) *The Origins of Angiosperms and Their Biological Consequences* (Cambridge Univ. Press, 1987).
- Crane, P. R. et al. *Am. J. Bot.* **74**, 1722–1736 (1987).
- Bino, R. J. et al. *Proc. Konink. Nederl. Akad. Wetensch. C* **87**, 1–24 (1984).
- van der Pijl, L. *Ann. Bogor.* **1**, 77–99 (1953).
- McWilliam, J. R. *Bot. Gaz.* **120**, 109–117 (1955).
- Kato, M. *Am. J. Bot.* **80**, 924–927 (1993).

Growth rings in Mesozoic birds

SIR — New discoveries of dinosaurs, together with cladistic procedures, have provided further insights into the morphological transition from nonavian dinosaurs to modern birds. Until now, however, the bone histology associated with this evolutionary transition has remained undocumented. Here we report the occurrence of growth rings in the Cretaceous birds *Patagopteryx deferrariisi* and Enan-

tiornithes, which are phylogenetically intermediate between nonavian dinosaurs and ornithurine birds¹, thus documenting for the first time this histological pattern among birds.

It has long been recognized that the bone tissue of nonavian dinosaurs is richly vascularized and endowed with dense haversian bone, a condition which has been cited as evidence of endothermy². But growth rings or lines of arrested growth (LAGs) (typical in ectotherms) in isolated bones^{3–5} and in the growth series of limbs of nonavian dinosaurs^{6,7} have been observed. In contrast, histological

studies of bones of extant birds and the Cretaceous ornithurine *Hesperornis*⁸ have never revealed growth rings interrupting the deposition of fibro-lamellar bone^{8–10}. This report of LAGs in Mesozoic bird bones suggests that the evolutionary history of avian endothermic homeothermy is more complex than previously thought.

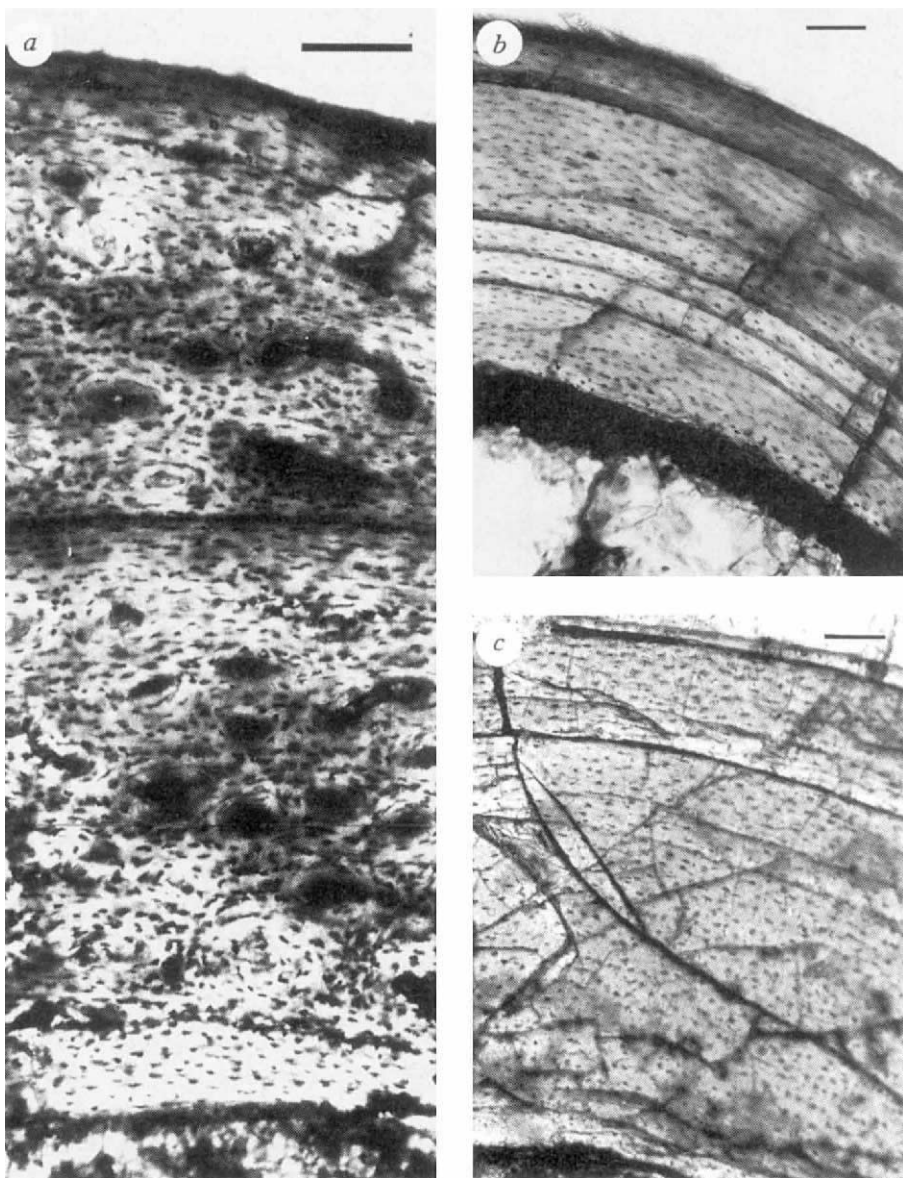
Patagopteryx was a hen-sized flightless bird found in the Coniacian-Santonian Rio Colorado Formation of southern Argentina¹¹. Thin sections¹² of the femur (MACN-N-03) reveal a distinct compact bone wall enclosing a medullary cavity devoid of cancellous bone tissue. The compacta, consisting of fibro-lamellar bone tissue (formed by rapid osteogenesis), is interrupted by a LAG marking a pause in bone deposition (see figure). Just internal to the LAG is a narrow band of lamellated tissue. The fibro-lamellar pattern of osteogenesis resumed after deposition of the LAG.

The Enantiornithes comprise a diverse group of volant birds distributed throughout the world¹³. The two enantiornithine femora (MACN-S-01; PVL-4273) studied are from the Maastrichtian Lecho Formation of northwestern Argentina¹⁴. The femora have a thin-walled, parallel-fibred compacta (formed by slow osteogenesis) enclosing a large medullary cavity. Deposition of this poorly vascularized compacta is interrupted by the occurrence of distinct LAGs (see figure); four in MACN-S-01 and five in PVL-4273.

The lack of fibro-lamellar bone in the enantiornithines is noteworthy. It is unlikely that this bone was completely resorbed during medullary expansion, as the remaining tissue should then resemble peripheral late bone, which it does not. The occurrence of widely spaced LAGs is unlike the closely spaced rest lines seen in the peripheral regions of bones of mature birds^{15–17} and mammals. Furthermore, resorption leaving only the peripherally lamellated late bone has never been reported.

The single LAG in *Patagopteryx* could result from a traumatic event, although the discovery of multiple LAGs in the enantiornithine birds strongly suggests that these interruptions in growth are genetic, reflecting a true zonal pattern. This pattern indicates that *Patagopteryx* and the enantiornithines grew more slowly than extant birds. Cyclical bone deposition occurs annually in various reptiles. Because no other cycle produces such a pattern⁴, we assume that LAGs in these birds formed annually, implying that they required more than 1 year to reach maximum body size.

Primary bone microstructure is the most reliable type of evidence used to infer physiology of extinct animals. Cyclical bone deposition in *Patagopteryx* and the Enantiornithes suggests that these birds differed physiologically from their



a, Transverse section¹² of the right femur the holotype (MACN-N-03) of *P. deferrariisi*. Note the single LAG and narrow region of lamellated bone interrupting the deposition of fibro-lamellar bone in the compacta. Several predominantly longitudinally oriented primary osteons are embedded in the woven bone matrix of the fibro-lamellar bone. The medullary cavity is lined by a narrow band of endosteally deposited lamellated bone tissue. b, c, Transverse sections of enantiornithine femora showing the poorly vascularized compacta interrupted by several LAGs. Large globular osteocytes with numerous canaliculi branching from them, as well as a smaller type with fewer canaliculi can be distinguished. b, Specimen PVL-4273 showing five LAGs. c, Specimen MACN-S-01 showing four LAGs. Scale bars, 100 μ m.

living relatives and were not fully homeothermic. A modern avian endothermic homeothermy has been suggested for the earliest bird *Archaeopteryx* on the basis of theoretical predictions^{18,19} and allometric studies²⁰. Similarly, inferences of flight capabilities have been the basis for assuming an endothermic physiology in the Early Cretaceous bird *Sinornis*²¹. Our study, documenting the presence of LAGs in more derived birds such as *Patagopteryx* and Enantiornithes, suggests that physiological interpretations of basal birds and nonavian dinosaurs requires serious reconsideration. Further-

more, it suggests that during the transition from nonavian dinosaurs to birds, classic endothermic homeothermy evolved after, not before²², the acquisition of feathers.

Anusuya Chinsamy, Luis M. Chiappe*
Peter Dodson

*Department of Animal Biology,
School of Veterinary Medicine,
University of Pennsylvania,
Philadelphia,
Pennsylvania 19104-6045, USA*

*Department of Vertebrate Paleontology,
American Museum of Natural History, New York,
10024, USA

- Chiappe, L. M. in *3rd Symp. Soc. Avian Paleont. Evol.* (ed. Peters, D. S.) (Courier Forschungsinstitut-Senckenberg, in the press).
- Bakker, R. T. *The dinosaur Heresies* (Morrow, New York, 1986).
- Reid, R. E. H. *Nature* **292**, 49–51 (1981).
- Reid, R. E. H. *Mod. Geol.* **15**, 19–48 (1990).
- Ricqlès, de A. *Acta palaeont. pol.* **28** (1–2), 225–232 (1983).
- Chinsamy, A. *Palaeont. afr.* **27**, 77–82 (1990).
- Varricchio, D. J. *vert. Palaeont.* **13**(1), 99–104 (1993).
- Houde, P. W. *Auk* **104**, 125–129 (1987).
- Enlow D. H. & Brown S. O. *Tex. J. Sci.* **9**(2), 186–214 (1957).
- Chinsamy, A. in *3rd Symp. Soc. Avian Paleont. Evol.* (ed. Peters, D. S.) (Courier Forschungsinstitut-Senckenberg, in the press).
- Alvarenga, H. M. F. & Bonaparte, J. F. in *Pap. Avian Paleont. Honoring Pierce Brodkorb* (ed. Campbell, K. E.) 51–64 (Nat. Hist. Mus. Los Angeles, 1992).
- Chinsamy, A. & Raath M. A. *Palaeont. afr.* **29**, 39–44 (1992).
- Chiappe, L. M. *J. vert. Pal.* **12**(3): 344–350 (1992).
- Walker, C. A. *Nature* **292**, 51–53 (1981).
- Soest Van, R. W. M. & Utrecht, W. L. *Bijdragen tot de Dierkunde* **41**(1), 61–66 (1971).
- Lewis, J. C. *J. Wildl. Mgmt.* **43**(1), 269–271 (1979).
- Nelson, R. C. & Bookhout, T. A. *J. Wildl. Mgmt.* **44**(2), 518–521 (1980).
- Bock, W. J. in *The Beginnings of Birds* (eds Hecht, M. K. Ostrom, J. H., Viohl, G. & Wellnhofer, P.) 199–207 (Proc. Int. Archaeopteryx Conf. Freunde des Jura-Museums, Eichstätt, 1985).
- Reig, P. J. in *The Beginnings of Birds* (eds Hecht, M. K. Ostrom, J. H., Viohl, G. & Wellnhofer, P.) 199–207 (Proc. Int. Archaeopteryx Conf. Freunde des Jura-Museums, Eichstätt, 1985).
- Houck, M. A., Gauthier, J. A. & Strauss, R. E. *Science* **247**, 195–198 (1990).
- Sereno, P. & Rao, C. *Science* **255**, 845–848 (1992).
- Ostrom, J. H. *Q. Rev. Biol.* **49**, 27–47 (1974).

Cerebral asymmetry and handedness

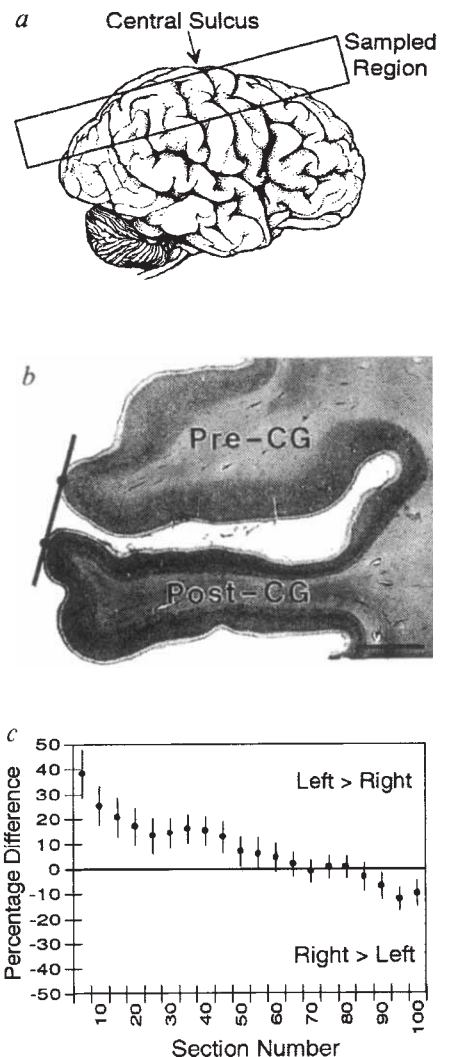
SIR — It has long been supposed that differences in the structure of the two sides of the brain underlie lateralized behaviour, perhaps the most compelling example being the preference in nine out of 10 of us for the right hand to perform complex manual tasks¹. Here we report measurements of the human central sulcus suggesting that the region of the brain that governs the upper extremity does indeed differ in the two hemispheres.

We examined the extent of cortical surface within the dorsolateral portion of the central sulcus bilaterally in the brains of 22 adults who died of non-neurological conditions (table). This sulcus is formed by the pre- and postcentral gyri, which contain the primary motor and somatosensory representations of the contralateral body^{2,3}. We cut the brains horizontally, and excised blocks containing the pre- and post-central gyri from both sides (a in the figure). Although the locations of the elements of the contralateral sensorimotor map along the central sulcus are quite variable, the region we sampled should include the representation of the upper extremity in most individuals^{2,3}. The tissue was cryoprotected by immersion in buffered sucrose solutions, frozen on dry ice, and serially sectioned with a

sliding 200-µm-thick microtome.

We analysed every fifth section, thus providing one detailed measurement from each millimetre of brain thickness. We confirmed the identity of the central sulcus microscopically by the characteristic laminar features of the pre- and post-central gyri⁴. The extent of the cortical surface within the sulcus was defined by a line drawn tangential to the crests of the gyri, traced, and measured (b in the figure). On average, the sulcal surface on the dorsolateral aspect of the cerebrum was greater in the left hemisphere than in the right (mean percentage difference, 7.2%). This asymmetry was most pronounced in the upper portion of the region sampled (c), and was present in 14 of the 22 brains examined.

These results suggest that humans have more cortical (and presumably subcortical) circuitry devoted to the representation



a, Location of the sampled region in a lateral view of the human brain; b, photomicrograph of a 200-µm-thick horizontal section through the central sulcus; c, average percentage differences of measurements from corresponding sections of the left and right hemispheres in the 22 brains examined. All sections were stained with cresyl violet acetate, coverslipped and coded so that the investigators did not know the hemisphere from which the section was taken. To provide a standard starting point, our analysis began with the first section in each hemisphere in which the pre- and post-central gyri were connected, and extended over the next 100 serial sections (measuring every fifth section). After projecting the section onto a digitizing tablet, a line was drawn tangentially to the crests of the pre- and post-central gyri (pre-CG and post-CG, respectively); the cortical surface between the tangent points (indicated by dots) was then traced and measured using morphometric software (Bioquant IV, R&M Biometrics, Inc.). The human central sulcus contains cytoarchitectonic fields 4 (primary motor cortex) and 3a, 3b, 1 and 2 (major subdivisions of primary somatosensory cortex)¹⁷. Scale bar b, 2 mm. In c, differences are $\{(L - R)/(L + R/2)\} \times 100$; $\pm$ s.e.m. A runs test for consecutive pairs of values indicated that this distribution was unlikely to have arisen by chance ($P < 0.0015$).

Age and brain weight of subjects examined			
	Number of subjects	Age (yr)	Brain weight (g)
Males	15	61.5 $\pm$ 12.6 (36–78)	1,401 $\pm$ 114 (1,200–1,650)
Females	7	54.9 $\pm$ 22.4 (17–76)	1,236 $\pm$ 109 (1,060–1,420)

Values are means $\pm$ s. d.; ranges in parentheses. All brains were acquired in conformity with Duke University Hospital guidelines and regulations. Determination of handedness in a postmortem study is problematic because this information is not included in the records of patients without neurological problems; in any event, this information cannot be considered authoritative. Based on the distribution of hand preference in the general population, 90% of our subjects should be right-handed¹.

of the right upper extremity than to the left. This suggestion accords with neuroanatomical studies of the spinal cord in man, which suggest that the corticospinal tracts are larger on the right side of the cervical enlargement^{5,6}. Many of these fibres originate in the left hemisphere (primarily those of the lateral corticospinal tract), and terminate in the right ventral horn of the cervical cord, where they innervate motor neurons governing the right upper extremity and related axial musculature. Moreover, electrophysiological studies of the primary motor map in non-human primates indicate that the somatotopic representation of the preferred paw is usually more extensive than that of the non-preferred paw⁷. Finally, volumetric measurements show that the right hand of right-handers is significantly larger than the left⁸. Based on what is known about trophic interactions between neurons and their peripheral targets⁹, the larger size of the right hand also implies a greater amount of neural circuitry devoted to the governance of this extremity.

This conclusion is supported by other lines of evidence linking extraordinary performance to increased cortical representation. Examples are the remarkable amount of cortical space devoted to the rhinarium of pigs^{10,11}, the lips of sheep and goats¹⁰, the tail of spider monkeys¹², the paws of raccoons^{13,14}, the whiskers of rats and mice¹⁵, and the rays of star-nosed moles¹⁶. We suggest that hand preference

in man is also instantiated by a greater amount of circuitry in corresponding neural centres. The gross measurements of the central sulcus that we have reported here argue for a detailed lateral comparison of the cytoarchitectonic areas along the entire human central sulcus.

**Leonard E. White, Greg Lucas
Ann Richards, Dale Purves**
Duke University Medical Center
Department of Neurobiology
Durham, North Carolina 27710, USA

- Coren, S. *The Left-Hander Syndrome: The Causes and Consequences of Left-Handedness* (Free Press, New York, 1992).
- Penfield, W. & Rasmussen, T. in *The Cerebral Cortex of Man* (MacMillan, New York, 1950).
- Woolsey, C. N. et al. *J. Neurosurg.* **51**, 476–506 (1979).
- Braak, H. *Architectonics of the Human Telencephalic Cortex* (Springer, New York, 1980).
- Yakovlev, P. & Rakic, P. *Trans. Am. Neurol. Ass.* **91**, 366–367 (1966).
- Nathan, P. W., Smith, M. C. & Deacon, P. *Brain* **113**, 303–324 (1990).
- Nudo, R. J. et al. *J. Neurosci.* **12**, 2918–2947 (1992).
- Purves, D., White, L. E. & Andrews, T. J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Purves, D. *Body and Brain: A Trophic Theory of Neural Connections* (Harvard University Press, Cambridge, 1988).
- Adrian, E. D. *Brain* **66**, 89–104 (1941).
- Woolsey, C. N. & Fairman, D. *Surgery* **19**, 684–702 (1946).
- Chang, H. T. et al. *Fedn. Proc.* **6**, 89 (1947).
- Welker, W. I. & Campos, G. B. *J. comp. Neurol.* **120**, 19–36 (1963).
- Johnson, J. I. in *Comparative Neurology of the Telencephalon* (ed. Ebessou, S.O.E.) 423–447 (Plenum, New York, 1980).
- Woolsey, T. A. & Van der Loos, H. *Brain Res.* **17**, 205–242 (1970).
- Catania, K. C. et al. *Nature* **364**, 493 (1993).
- Zilles, K. in *The Human Nervous System* (ed. Paxinos G.) 752–802 (Academic, San Diego, 1990).

Angiogenic activity of enzymes

SIR — We have purified thymidine phosphorylase (dThdPase), an enzyme involved in pyrimidine nucleoside metabolism, from human placenta¹. The activity of this enzyme has been reported to increase in several types of malignant

tumours in man^{1,2}, but little is known about its precise physiological functions. We have shown that 120 amino acids of human dThdPase are identical to the sequence of platelet-derived endothelial cell growth factor (PD-ECGF)³, and that

recombinant PD-ECGF has dThdPase activity^{4,5}. PD-ECGF was originally isolated as an endothelial cell mitogen from platelets⁶; it has chemotactic activity *in vitro* and angiogenic activity *in vivo*^{7,8}, but does not stimulate cell proliferation⁹. Here we show that dThdPase also has chemotactic and angiogenic activity.

The figure shows that bovine aortic endothelial (BAE) cells migrated in response to PD-ECGF and dThdPase in a manner that was dose-dependent. We took two different approaches to investigate whether dThdPase has angiogenic activity *in vivo*. First, we investigated the

effect of dThdPase on the developing vascular system of the chick chorioallantoic membrane (CAM assay) and found that recombinant PD-ECGF and purified dThdPase induce a dose-dependent angiogenic response. Second, we investigated the angiogenic activity of compounds in gelatin sponges implanted into a mouse subcutaneous pouch¹⁰. The haemoglobin content extracted from each sponge is a parameter of vascularization. The angiogenic response was evident at a dose of 10 µg basic fibroblast growth factor (bFGF) and dThdPase.

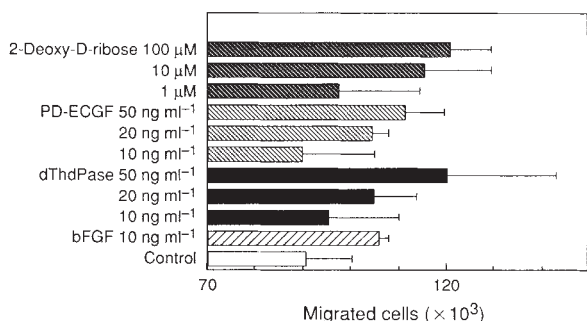
A competitive inhibitor of dThdPase, 6-amino-5-chlorouracil, almost completely inhibits the angiogenic effect of dThdPase, but does not significantly inhibit that of bFGF. These findings suggest that dThdPase itself is responsible for the angiogenic activity, and that the dThdPase activity is indispensable for the angiogenic effect of dThdPase.

We thus examined the angiogenic activity of the degradation products of thymidine by dThdPase as well as 2-deoxy-D-ribose, a dephosphorylated product derived from 2-deoxy-D-ribose-1-phosphate. Angiogenesis was induced in 72% of the eggs by 6.7 ng (50 pmol) 2-deoxy-D-ribose. Thymidine, thymine, 2-deoxy-D-ribose-1-phosphate and D-ribose do not have significant angiogenic activity at the same dose as that of 2-deoxy-D-ribose.

In conclusion, PD-ECGF, dThdPase and 2-deoxy-D-ribose have chemotactic activity *in vitro* and angiogenic activity *in vivo*. PD-ECGF does not induce angiogenesis by direct stimulation of endothelial cell proliferation⁹. The enzymatic product(s) may stimulate the chemotaxis of endothelial and possibly other cells, causing angiogenesis: 2-deoxy-D-ribose may be one such product responsible for angiogenesis by dThdPase.

**Misako Haraguchi, Kazutaka Miyadera,
Katsuo Uemura, Tomoyuki Sumizawa,
Tatsuhiko Furukawa, Kazutaka Yamada,
Shin-ichi Akiyama**

Institute for Cancer Research and First
Department of Surgery,
Kagoshima University,
Sakuragaoka, Kagoshima 890, Japan
Yuji Yamada
Taiho Pharmaceutical Co Ltd,
Saitama 357, Japan



Chemotactic activity of dThdPase. The number of BAE cells which migrate towards dThdPase, bFGF, PD-ECGF and 2-deoxy-D-ribose is shown. Values are the average ± s.d. of three determinations. The chemotaxis assay using BAE cells was performed as described⁷. Collagen coating of the upper wells was necessary for BAE cell adhesion. BAE cells (1 × 10⁶ cells per well) were added to the upper wells and the compounds at the indicated concentrations in modified Eagles Medium supplemented with 1% fetal calf serum were added to the lower wells. For the control, the medium supplemented with 1% serum was added. The number of cells migrating to the lower surface of the filter was determined.

- Yoshimura, A. et al. *Biochim. biophys. Acta* **1034**, 107–113 (1990).
- Pauly, J. L. et al. *J. nat. Cancer Inst.* **58**, 1587–1590 (1977).
- Furukawa, T. et al. *Nature* **356**, 668 (1992).
- Usuki, K. et al. *Biochem. biophys. Res. Commun.* **184**, 1317–1323 (1992).
- Sumizawa, T. et al. *J. Biochem.* (Tokyo) **114**, 9–14 (1993).
- Miyazono, K. et al. *J. Biol. Chem.* **262**, 4098 (1987).
- Ishikawa, F. et al. *Nature* **338**, 557–561 (1989).
- Folkman, J. & Shing, Y. *J. Biol. Chem.* **267**, 10931–10934 (1992).
- Moghaddam, A. & Bicknell, R. *Biochemistry* **31**, 12141–12146 (1992).
- Presenti, E., Sola, F., Mongelli, N., Grandi, M. & Spreafico, F. *Br. J. Cancer* **66**, 367–372 (1992).

Buy now, pay later

Tom Kirkwood

Human Longevity. By David W.E. Smith. Oxford University Press: 1993. Pp 175. £27.50. \$35.

"If ye would live long, choose well thy ancestors", advises the Good Book. And it is not a bad idea, apparently, to be female, Japanese, a non-smoker, a light drinker and to enjoy sex, though not all of these options are open to everyone.

The science of human longevity is a game for many players. First, there is the numbers game of demography: how long do we live, and how is this changing? Second, there is the zoo quest of comparative gerontology: how does our longevity compare with that of other species? Next come the doctors and nurses: what do we actually die of? Biochemists and molecular biologists are key players too: what causes the diseases and functional decrements of old age? Sociologists join the team when behavioural and social aspects of longevity are addressed: who lives longest?

David Smith's admirable book tackles each of these issues. For such a large topic, the book is surprising short, just 130 pages of main text. Brevity is both a strength and a weakness. The strength is that the book is easy to read as it skips lightly through a broad array of material. The weakness is that the appetite is whetted, not sated, and inevitably there are gaps.

The demography of lifespan is fascinating. We all know that average lifespan has increased, doubling over the past few hundred years in the developed countries of the world. But we need reminding that most of the change is due to reduction in child mortality. Simple arithmetic shows us that if nearly half of children die before their fifth birthday, and if child mortality is reduced, then the average lifespan jumps dramatically. In seventeenth-century England, for instance, life expectancy at birth was around 35 years. But for a 30-year-old who had survived the dangers of early life, life expectancy was still another 30 years. There is little evidence that maximum lifespan has grown, except in the obvious sense that the longest recorded lifespan can only get longer.

Along with the change in life expectancy in developed countries, there have been large changes in the causes of death. Infectious diseases no longer head the list; cardiovascular disease and cancer have taken their place. A difficult question, touched on but unresolved, is the extent to which old age itself may in the absence of disease or trauma be the cause of death. Death certificates, especially for the very old, are unreliable; many are just plain wrong.

Life insurance costs more for men than women because male mortality is higher at all ages. Females live about 6 years longer than males. There are countries where the difference is small, or even reversed, but these are countries where sex inequality in society is extreme. In Bangladesh, for instance, Smith reports that boys under five are given 16 per cent more food than girls, the difference increasing in times of famine. Bangladesh has short life expectancies at birth both for men (54.9 years) and women (54.7 years), but men live marginally longer. There is an interesting point here. The only known way to extend lifespan greatly in mammals is by severe dietary restriction, which in rodents can increase longevity by 30 per cent or more. Elsewhere, Smith discusses research into the generality of this phenomenon, including trials in primates. The evidence from Bangladesh is not encouraging.

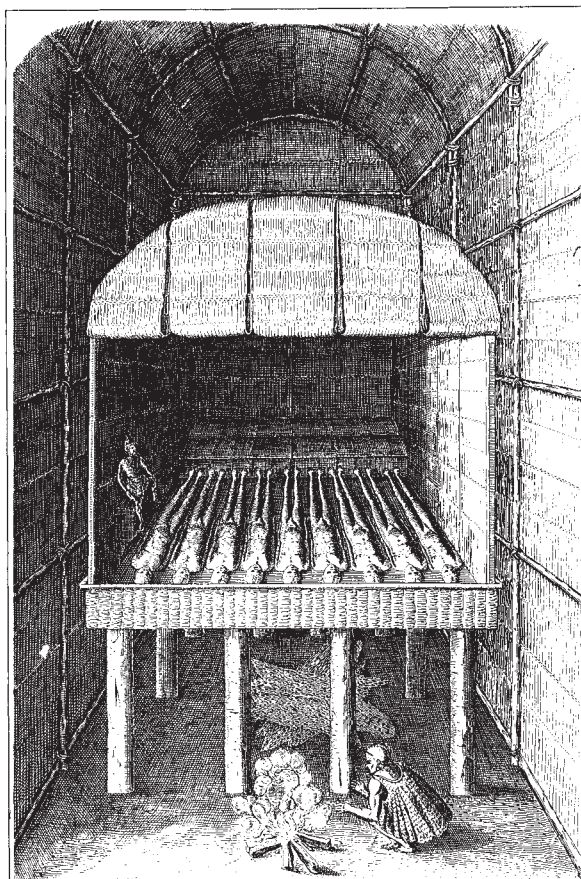
The biological basis of sex differences in longevity is interesting and so far unresolved. Smith discusses in some detail the theory that sex chromosomes have a role to play, but dwells rather less on endocrine factors. Hormones are important modulators of physiology and behaviour and there is evidence, although scant, that castration increases male longevity. Castration does not do a lot for the chromosomes, but it does have a striking effect on hormones. Perhaps unsurprisingly, there are few takers for this route to rejuvenation.

Choosing the right ancestors is a big help. They should be rich because wealth brings health. They should also have the right genes. Genes that have life-shortening effects should be avoided, such as those that predispose to cancer or early-onset Alzheimer's disease. Of deeper interest, however, are genes that modulate the rate of ageing. It is not yet known what these genes are, but strong hints are emerging from

evolutionary theories of ageing and longevity. Some will be genes for maintenance, such as the antioxidant defences of cells. Maintenance is good but may be costly, and a degree of tuning through natural selection is to be expected. Others may be 'buy now, pay later' genes which benefit the young but cripple the old. Smith reviews the main evolutionary theories and asks why humans have evolved a lifespan so much longer than that of other mammals.

The science of human longevity can only grow, as more people turn their attention to this age-old problem. We urgently need to address the fundamental research that some day, it is hoped, will equip us to improve the quality of later life. Smith has done a fine job of introducing the topic to a broad potential readership. □

Tom Kirkwood is in the School of Biological Sciences and Department of Geriatric Medicine, University of Manchester, 3.239 Stopford Building, Oxford Road, Manchester M13 9PT, UK.



ON THE RACK — burial practices among American Indians. Taken from the frontispiece of *Disease and Demography in the Americas* edited by J. W. Verano and D. H. Ubelaker, a collection of 23 regional essays on the biological and cultural impacts of European colonization of the New World. Now published in paperback by Smithsonian University Press, \$29.95.

Corpuscular science

David Knight

Robert Boyle Reconsidered. Edited by Michael Hunter. Cambridge University Press: 1994. Pp. 231. £33, \$44.95.

"Divulging of Useful Truths in Physick": The Medical Agenda of Robert Boyle. By Barbara Beigun Kaplan. John Hopkins University Press: 1993. Pp. 216. \$40, £33.

As "Father of Chemistry and Uncle of the Earl of Cork", Robert Boyle's name is better known than his works. He is nevertheless remembered for his law of gas pressure and volume and often praised for his definition of a chemical element, which was not new and which he did not consider to be especially useful. He saw the world as made up of 'corpuscles' (a word revived in 1897 by J. J. Thomson for the components of cathode rays) that were all composed of the same matter rather than irreducible elements. His *Sceptical Chymist* makes heavy reading; and indeed his style, full of qualifications, parentheses and digressions, has been to later readers an obstacle to understanding his thought.

Boyle's rhetoric is one of the topics addressed in the essays in *Robert Boyle Reconsidered*, which began in a conference to mark the three-hundredth anniversary of Boyle's death in 1691. If we see Boyle as self-consciously a gentleman, anxious to avoid controversy or confrontation, reluctant to name names, growing up and living in a period of religious and civil wars, and keen to emphasize what united rather than divided people, then his tortuous style makes sense. Gentlemen did not put themselves forward as authors but preserved their amateur status: and so Boyle's writings were often not published for years after they had been written, and then only after a little urging. Prolixity and qualification were features of Boyle's diplomacy in making allies but also in distancing himself from them.

These studies are based on close readings of Boyle's papers at the Royal Society of London, where there is a huge archive of around 20,000 pages. The volume is an indication of the state of Boyle scholarship, with a new edition of his works imminent. Just as our view of Newton has been transformed by work on his papers, so these essays open up new perspectives on Boyle. As we might expect, they place him more closely in the context of his own

day: a most important time, in which the first lasting scientific institutions were formed and the 'new philosophy' gained wide acceptance. Boyle's position in the scientific world, to which he brought prestige, is familiar, and we see him here relating interestingly to Robert Hooke, his social inferior; but some may be surprised by Boyle's concerns with what we would call philosophy as well as with science, his deep theological commitments and his abiding interest in alchemy. In "*Divulging of Useful Truths in Physick*" Barbara Kaplan explores the medical thinking of this known hypochondriac. All this fits well with someone of Boyle's background and interests, and these new works bring Boyle to life as a person engaged in natural philosophy; science is always done by people, after all.

Boyle has often been thought of as a mechanist, hoping to explain everything

but that God's purposes were not (as Cartesians supposed) completely inaccessible to us: we know that the eye is for seeing, and final causes were an essential complement to mechanism. They also, by confuting the Epicureans' world of chance, reinforced religion.

Boyle opposed the idea that nature was a kind of vicegerent or demiurge, put in charge of the material realm: God created and supported the world directly, just as He dealt directly with the believer. The world was a complex place, not always to be understood in mathematical terms: Boyle's science after all was chemistry. Three essays deal with his chemical views, emphasizing his devotion to experiment, his opposition to mere empiricism and his alchemy. His corpuscles are normally traced back to the atoms of Democritus, Epicurus and Lucretius; but it is argued here that they have equally close affinities with those found in the writings of 'Gabriel' and later alchemists. Boyle read widely in alchemy, and his corpuscular philosophy meant that alchemy was not impossible; his chemistry had close links with what we might consider occult science. The middle ways he tried to follow in science, politics and religion were strait and narrow.

Anybody interested in the 'Scientific Revolution' or in how great scientists' minds work will find Hunter's book fascinating. The contributors have made their pieces accessible and the essays go well together. There is mention of Boyle and medicine.

As Kaplan shows, Boyle had no medical training, although he did work closely with some doctors and was regarded as an expert on the eye. He lived at a time of professional crisis when the book-learning on which physicians' prestige depended was being undermined. He insisted on the importance of experiment and science in medicine, as well as reading — in the tradition of William Harvey, but with corpuscular explanations where possible. He worked on respiration and blood transfusion, making Descartes' explanation of the heartbeat seem less plausible than Harvey's. But his chemistry could not really provide testable medical explanations and his medical legacy was not a very rich one. Kaplan's Boyle is recognizably the same as the Boyle of Hunter's volume, and Kaplan shows how one can cast light on scientists through their more peripheral studies. We see Boyle no longer overshadowed by Newton, and get a more just view of the science of his day. □

David Knight is in the Department of Philosophy, University of Durham, 50, Old Elvet Street, Durham DH1 3HN, UK.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Robert Boyle (1627–91)

in terms of matter and motion, and making great use of corpuscles (often in the form of 'effluvia') to explain anything and everything. His famous passage on the excellency and grounds of the mechanical hypothesis seems to confirm this, and his writings abound with the idea that mechanical explanations are the only genuine kind. But there are other messages there too; and five of the twelve essays in Hunter's volume are devoted to Boyle's thoughts about cosmic qualities, the limits of human reason, final causes and design, and miracles. He was well known for his piety and his scruples in matters of religion, positions which prevented him from taking the oaths required of a president of the Royal Society (he declined the office). Boyle believed that there were things we could never know,

Mary Evans Picture Library

Stellar beginnings

André Maeder

The Physics of Stars. By A. C. Philips. Wiley: 1994. Pp. 208. £42.50, \$68.50 (hbk); £17.95, \$28.95 (pbk).

FAR-reaching consequences in stellar physics can be derived from first principles. The beautiful and straightforward connections between fundamental physics and the properties of stars are shown well in this book. It is pitched at undergraduates and successfully provides short accounts of selected topics. The author appreciates that the motivation to study physics is enhanced by astrophysical applications, while a proper understanding of how stars work in turn requires a clear and sound basis in physics.

After a good introductory chapter on some important astrophysical concepts, the author focuses on the properties of matter and radiation. He derives the various equations of state that have large consequences for stars, and provides a brief account of the main processes of heat transfer and the basic physics of nuclear processes, giving an overview of the main nuclear reactions during stellar evolution. On treating the applications to stellar structure, he emphasizes simple models and minimum and maximum stellar masses. He then makes the wise choice of jumping to the end-points of stellar evolution (white dwarfs, neutron stars and black holes), while showing how basic principles determine their main features.

Throughout there is a proper balance between text and equations. The sequence of arguments and the logical steps in the mathematics are developed with care so that they can be easily followed by a reader with a basic knowledge of physics. My only criticism is the lack of good, relevant figures.

Problems are posed at the end of each chapter, and hints at the answers are given at the back of the book. Many such books contain problems so difficult that nobody even tries to solve them; the ones here are interesting and easy enough to stimulate rather than discourage students. □

André Maeder is at the Observatoire de Genève, CH des Maillettes 51, CH-1290 Sauverny, Switzerland.

■ Also just published: **Astrophysical Jets** edited by D. Burgarella, M. Livio and C. O'Dea. Proceedings of a meeting held in Baltimore in May 1992. CUP, £35, \$59.95; **The Nature of Compact Objects in Active Galactic Nuclei** edited by A. Robinson and R. J. Terlevich. Proceedings of the 23rd Herstmonceux conference, Cambridge, July 1992. CUP, £40, \$59.95; and **Solar and Planetary Dynamos** edited by M. R. E. Proctor, P. C. Matthews and A. M. Rucklidge. Proceedings of a NATO meeting, Cambridge, September 1992. CUP, £35, \$49.95.

Big bodies, big brains



MEAVE Leakey and Alan Walker examining the cranium of the *Homo erectus* boy found during the 1984 field season at Nariokotome, Kenya. The skeleton itself is the subject of the book reviewed below, while this picture is taken from *Ancestors: In Search of Human Origins* written by the palaeoanthropologist Donald Johanson together with Lenora Johanson and Blake Edgar (Villard Books, \$27.50). Immaculately illustrated and produced, the volume was released to coincide with a recent three-part NOVA television series — hosted by Donald Johanson — on human evolution.

Old boy

Bernard Wood

The Nariokotome Homo Erectus Skeleton. Edited by Alan Walker and Richard Leakey. Springer: 1993. Pp. 457. DM248.

EVOLUTIONARY history can often be retrieved by comparing closely related living forms. But modern humans differ substantially in both morphology and behaviour from their closest living relatives, the African apes, the chimpanzee and the gorilla. We are the sole living representative of our adaptive plateau, so details of our evolutionary history have to be sought from the palaeontological record.

Fossil evidence usually comes in dribs and drabs, a tooth here and a fragment of limb bone there. The challenge to palaeontologists is to use these data to generate testable hypotheses about human evolution. There is always the possibility that limbs may be matched with the wrong body and teeth with the wrong head, thus creating combinations that did not exist in nature. Now and again, however, much better preserved fossils are recovered. The main value of a single skeleton is that it can confirm or refute links between cranial and post-cranial remains as well as provide more reliable and otherwise unobtainable evi-

dence about relationships between body shape and size.

The early hominid skeleton KNM-WT15000 (referred to here as WT15K) is unrivalled among its contemporaries in its completeness. The first evidence of it, a small piece of cranial vault, was discovered by the veteran fossil-hunter Kamoya Kimeu in 1984 beside a tributary of the Nariokotome sand river on the west side of Lake Turkana in northern Kenya. Its recovery took five seasons of excavation and the removal, by hand, of 1,500 cubic metres of earth. Little of the feet, hands and forearms was recovered and there are no neck vertebrae, but, these deficiencies apart, the skeleton is remarkably complete. The geological reconstruction suggests that the skeleton was dismembered by decomposition and the components buried naturally in a swamp next to a grassy floodplain. Its geological age falls between 1.51 and 1.56 million years old. The cause of death is unknown but there is no evidence of violence: the only pathological sign is a small abscess cavity in the lower jaw.

Alan Walker and Richard Leakey, the editors of this volume, were both directly involved in unearthing and reconstructing the skeleton. The book is in two parts, the first describing the skeleton and its context and the second comprising the reports of eight analytical studies. It finishes with a summary by Alan Walker that puts the

work in perspective.

The skeleton belonged to an 11–15-year-old boy about 160 cm tall and weighing some 40 kg. (Its predicted adult height and weight would have been 185 cm and 68 kg.) The body proportions, with long limbs and narrow pelvis, resemble those of modern humans living in tropical climates. Not long ago the immaturity of the specimen would have been regarded as a disadvantage, but the crowns and roots of the teeth, for example, have now proved to be fertile sources of information about comparative development.

In particular, B. Holly Smith's comparative analysis of the physiological age of the skeleton is exemplary, demonstrating a lack of the distinctive human pattern of growth suppression in late childhood followed by a compensatory growth spurt in adolescence. Another characteristic of modern humans is that the brain continues to grow at the fast fetal rate during the first year after birth. In this respect WT15K is like modern humans. It is puzzling, however, that the relationship between brain and body size in WT15K is little advanced over that of the australopithecines who preceded and continued to live alongside it.

A short review can do scant justice to the range, quality and importance of many of the studies included in the volume. Particularly noteworthy are the contributions on the rib cage, which is predominantly like that of modern humans, and the canal for the spinal cord, which is also similar except for a relative narrowing in the thoracic region; these areas have received little attention from hominid palaeontologists. It is a pity that there are no studies of diet or locomotion, but no doubt these are in progress.

Taxonomic analysis reveals that WT15K belongs to the species many refer to as early African *Homo erectus* and some as *H. ergaster*. The species dates back to 1.8 million years ago, perhaps even earlier. It is a creature that has reached an adaptive plateau substantially more like ours than that of the australopithecines. More importantly and intriguingly, this plateau differs from that of two other species groups presently included in the genus *Homo* as *H. habilis* and *H. rudolfensis*. We may have cast the net too wide for the genus *Homo*.

This volume shows how to put the biology into palaeobiology. Despite the poor quality of the photographs, it sets an admirably high standard that will be hard to match. The next challenge is to understand the factors that constrained brain size until much later in hominid evolution. □

Bernard Wood is in the Department of Human Anatomy and Cell Biology, The University of Liverpool, PO Box 147, Liverpool L69 3BX, UK.

Creatures deep and small

John D. Gage

Meiobenthology: The Microscopic Fauna in Aquatic Sediments. By Olav Giere. Springer: 1993. Pp. 328. DM128, SFr128, £50.50, \$89.

THE study of the bottom-dwelling, or benthic, organisms inhabiting aquatic sediments has progressively focused on ever-smaller classes. This reflects the increasing need to understand the role of soft sediments in organic remineralization and pollution in both marine and freshwater environment, where the activity of microorganisms is paramount. The habitat is dominated by the deep-sea bed where much of the metazoan fauna has become miniaturized. Also, the 40 per cent of all fresh water on Earth thought to be stored in deep groundwater systems provides a habitat for microscopic metazoans distributed among deep caves, surface waters and even shallow marine sediments.

Some of these organisms turn out to belong to new phyla such as the Kinorhyncha, or the very recently described Loricifera. Others have been assigned to major new and supposedly primitive taxa, such as the Archiannelida and the new crustacean subclass Cephalocarida, discovered by Howard Sanders in the 1950s. The recognition of an important subcommunity of benthic organisms that can be seen only with a microscope came with pioneering work by Adolf Remane in Germany and later by Delamare Deboutteville in France on what became known as the 'interstitial' fauna of sediments. The terms micro-, meio- and macrobenthos were coined by Mare in 1942 to describe the main size classes in the marine benthic community. But it was not until the past decade that Peter Schwinghamer showed that these empirical divisions correspond to marked abundance peaks in biomass size spectra of benthic animals from a wide variety of aquatic sediments.

These peaks may be determined by physical constraints of the sediment habitat. One peak corresponds to the size range normally considered as 'meiofauna' (about 125–250 µm; interstitial lifestyle). Other peaks correspond to the microfauna (generally microflora attached to sediment particles, about 1 µm) and macrofauna (usually less than 250 µm; occupy the sediment).

Giere's scholarly yet readable monograph on the meiobenthos will be welcomed by all aquatic benthic ecologists. It signals a maturity in the development of this still rather obscure branch of marine biology that will undoubtedly gain wider

interest, despite huge taxonomic problems. Its thorough and well-balanced treatment, and useful suggestions for further reading as well as text citations, will also make the study of these fascinating animals more accessible, especially to young aquatic scientists (and particularly to those in the many areas of the world where there is no information at all on the meiobenthos). The well-drawn illustrations and informative summaries of the phylogenetically varied but usually rather conservative body form demonstrate the powerful constraints imposed on life between sediment particles. Indeed, the sediment itself has been taken as of most importance in determining the meiobenthic habitat. More recently, studies using microelectrodes have revealed often sharply discontinuous boundaries in other conditions: for example, high oxygen concentrations are usually associated with biogenic microstructures such as irrigatory burrows of macrofauna. Furthermore, the discovery of meiobenthic metazoans (mainly nematodes) and ciliate protozoan microbenthos that require reducing conditions rich in hydrogen sulphide but apparently completely lacking in oxygen has generated controversy and raised many questions — some see this fauna as a relict from ancient anoxic environments. Although the biogeochemical implications of this sulphide fauna may have been underestimated, it would be unsafe to claim, as did a recent estimate, that 50 per cent of benthic energy flux may be caused by sulphate reduction in shallow marine sediment.

It has been difficult enough to uncover the overall energetic role of meiofauna, and much uncertainty remains. On the one hand, close interaction with microorganisms and detritus has suggested a well-defined and largely closed 'detrital trophic complex'. Giere has examined this assumption, however, and argues instead for strong mutual links between the meio- and macrofauna long evident from the large part of the meiofauna comprised of larval stages of macrobenthos.

What is clear is the pressing need for benthic ecologists to examine the benthic ecosystem in a more comprehensive and integrated fashion. □

John D. Gage is in the Dunstaffnage Marine Laboratory, PO Box 3, Oban, Argyll PA34 4AD, UK.

Spring Books

Next week's issue (24 March) contains *Nature's* Spring Books supplement, featuring, among others, Richard L. Gregory on Francis Crick's *The Astonishing Hypothesis*; Christopher Longuet-Higgins on Steven Pinker's *The Language Instinct*; McGeorge Bundy on James B. Conant; and Thomas Banchoff on the philosophy of holes.

Ground-based imaging of extrasolar planets using adaptive optics

J. R. P. Angel

Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

The detection of extrasolar planets by direct imaging presents an extraordinary technical challenge. They must be identified against background light scattered from a star close by and about a billion times brighter. It has been supposed that a near-perfect space telescope would be required to avoid atmospheric blurring. But by using adaptive optics operating at fundamental performance limits, the new generation of large ground-based telescopes has the potential to detect planets orbiting nearby stars.

CURRENT search strategies for extrasolar planets are aimed at indirect detection by the small reaction motion of a star. To date no positive detection has been made, except for the Earth-mass objects orbiting a neutron star¹. Direct imaging could provide unambiguous and immediate detection, and even physical information through photometry or low-resolution spectroscopy. Observations in the thermal infrared have the extraordinary potential for detecting and determining atmospheric composition of Earth-like planets out to several parsecs distance², but large cooled telescopes in space are required³. Unfortunately the Hubble telescope will probably lack the sensitivity for optical detection of an extrasolar planet, even of the size and orbital radius of Jupiter, because of the relatively small aperture combined with static, small-scale wavefront errors⁴.

Ground-based telescopes have also been inadequate, because of atmospheric blurring, but now the technique of adaptive optics has been developed to correct atmospheric errors before light is brought to a focus. Planet detection has especially favourable conditions for correction, because of the host star's strong signal. But present and planned adaptive systems for astronomy are not designed to realize this potential, because wavefront measurements with the required high precision are not usually possible.

Here I explore the fundamental performance limits of adaptive optics to sense and correct the atmospheric distortion of a bright unresolved star, and describe a new type of system to reach this limit. This provides for phase measurements and correction to a higher accuracy and speed than has been previously considered necessary, even for military applications, and for correction of intensity fluctuations which for other purposes are insignificant. A few hours with a 6.5-m telescope corrected with this system would be sufficient time to detect, at the 5 standard deviation (s.d.) level, planets with Jupiter's size and orbit at nearby stars such as τ Ceti or π^3 Orionis, or an Earth-like planet of the singularly close binary star α Centauri.

This Article first introduces the method of adaptive optics, then examines the intensity and extent of scattered light in the halo surrounding an adaptively corrected stellar image, against which the planet must be detected. Critical factors are the amplitude of the residual phase errors and the smallest spatial scale of correction. I show how the interferometric metrology method used for precision optical testing can be adapted to make a practical stellar wavefront sensor, the accuracy of which is fundamentally limited by noise due to the random arrival of photons. The device is also optimal for detecting very small phase aberrations that develop as turbulence alters the wavefront between the successive corrections. An expression is derived for the limiting accuracy of phase error measurement, set only by the brightness of the star and the characteristics of the prevailing atmospheric turbulence.

With phase errors corrected to this limit, it is found that a previously insignificant effect from intensity modulation (atmospheric scintillation) will now dominate the halo. With this error also corrected, the star and planet images will be formed as stable Airy patterns corresponding to the full telescope aperture. The small residual error results in an underlying faint diffuse halo of scattered starlight several arc seconds in diameter, and showing speckle structure from interference that changes after each cycle of adaptive correction. This structure, along with photon noise, limits the sensitivity in any given exposure. Finally, I project the optimized design parameters and sensitivity for 6.5-m and 8.4-m binocular telescopes under construction. For observations at 0.8 μm wavelength, the incoming wavefront must be measured and corrected over $\geq 10,000$ spatial pixels at a frequency of ≥ 2 kHz to obtain the performance summarized above.

Residual aberration after adaptive correction

In this section I briefly review the principles of adaptive optics and examine how the halo of scattered light around a star image depends on the correction. Plane light waves from an unresolved star become distorted on passing through regions of variable refractive index in the atmosphere. Adaptive correction of the resulting blurred images involves repeated measurements of the instantaneous wavefront phase aberration, and immediate compensation in real time by the superposition of an equal and opposite aberration, before the light is brought to a focus at the detector. The compensation is typically introduced in a focal-plane instrument by reflection on a small mirror, the shape of which can be rapidly deformed. If the light from the bright star is corrected in this way, any planet appearing within a small surrounding field of a few arc seconds, called the isoplanatic patch, will be similarly corrected⁵. Given good correction, diffraction-limited images will be formed of both star and planet, of angular width λ_m/D , the diffraction limit for a telescope of diameter D at imaging wavelength λ_m . However, even very small residual errors can add a halo around the star that is still brighter than the planet.

It is in principle possible to build a correction device that will alter the wavefront to any prescription as rapidly as desired. In the discussion that follows, I will assume that errors from the device are held at a level negligible compared to the fundamental measurement error. The width and central concentration of the halo then depends on the strength and distribution of the residual measurement errors among different spatial frequencies. Suppose that wavefront phase measurements are obtained as averages over small areas (pixels) whose linear dimension, projected on the telescope pupil, is Δx and that the measurement technique results in errors uncorrelated between adjacent pixels. Under these conditions, after correction the halo caused by the

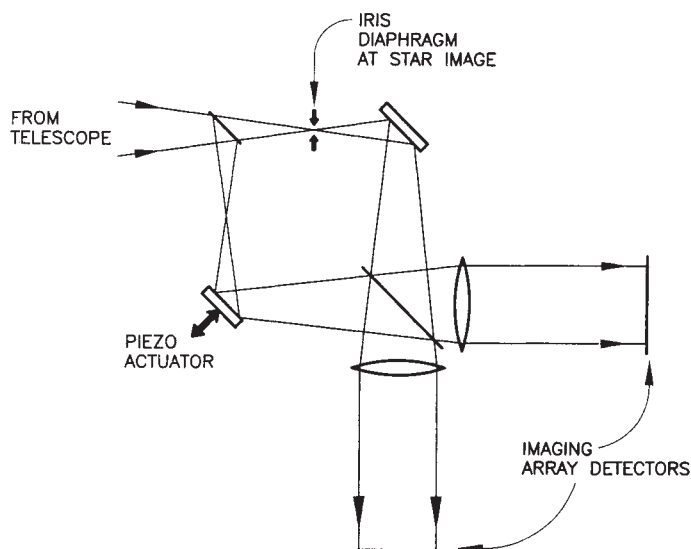


FIG. 1 Mach-Zehnder interferometer adapted for use as a phase-shifting stellar wavefront sensor.

measurement errors is not strongly peaked, with central intensity reduced from that of the star by approximately σ_m^2/n , where n is the total number of pixels and σ_m^2 is equal to the residual mean square phase error in radians² at the imaging wavelength. (This follows directly from considering a point in the image where, in the absence of errors, amplitudes summed over the pupil add to zero. When errors are included, the n phase residuals add as a random walk.) The fraction, f , of the total starlight energy scattered into a halo as a result of small residual phase errors is given by the Maréchal approximation⁶, $f = \sigma_m^2$. It follows that the halo width is larger than the telescope diffraction width by a factor $\sqrt{n} \approx D/\Delta x$, and is thus approximately equal to $\lambda_m/\Delta x$. We will find below that in practice after accurate, fine-scale correction this width is considerably larger than natural atmospheric blurring, and is likely to extend beyond any detectable planet. Scattering by dust and dirt on the mirrors may be comparable in total energy to that from phase errors, but because the scale is much smaller, the corresponding halo will be correspondingly larger, with negligible surface brightness.

A critical factor in determining limiting sensitivity will be the gain, G , of the corrected star image, defined as the star's peak intensity relative to that of the scattered halo. It follows from the above that this is given by

$$G = 1/\sigma_m^2 (D/\Delta x)^2 \quad (1)$$

The highest gain is thus given by adaptive correction that minimizes the wavefront error and has the finest scale of correction of the largest-diameter telescope. In practice the ratio of star brightness to planet brightness is expected to be of the order of 100 times larger than even the highest achievable value of G , and thus the halo must be measured with high signal-to-noise ratio (S/N) through long exposures.

The wavefront sensor

The most critical element of an adaptive optics system to search for planets is the wavefront sensor. The sensors commonly used for adaptive optics, based on wavefront slope measurements⁵, do not take advantage of the spatial coherence of the stellar wavefront⁷. This is done in the most accurate measurements in the laboratory, which employ direct interference with an unaberrated coherent laser reference. Fringe patterns recorded with predetermined, stepped phase shifts between the aberrated wavefront and reference beam allow for measurement of very small phase differences, with relative freedom from calibration errors⁸.

Here I describe a new type of sensor based on the Zernike test⁹, that adapts these techniques to measure atmospheric aberration

(Fig. 1). The light converging to a star image passes through a beamsplitter into the two arms of a Mach-Zehnder interferometer⁶. A small aperture acting as a spatial filter transmits a reference beam that is almost perfectly spherical. The direct beam is reflected unobstructed off a piezo-driven mirror which is repeatedly stepped through a full wave of path difference. The beams are recombined with zero average path difference at a second beamsplitter that yields two outputs with white light interference fringes. These are sensed with array detectors at images of the telescope pupil. The phase difference between the two wavefronts, averaged over each square pixel of the array, is calculated from fringes recorded at each setting of the piezo-driven mirror. The first experimental stellar interferometer of this type has recently been built and tested at the 2.3-m telescope of Steward Observatory by Colucci¹⁰. Phase maps were obtained of a bright star wavefront, from interferograms recorded every 3 ms when a bright speckle happened to fall on the reference aperture.

The full power of this method will be realized in our application, when a servo control loop is closed about the interferometer measurements to continually correct the wavefront. Then there is always a strong reference signal. It is supposed that initially the wavefront has been corrected to a mean square error of 1 rad² or less, using either conventional methods or a bootstrap method in which the spatial filter is progressively reduced in diameter. At this point the spatial filter is set at the diameter of the first dark Airy ring and the servo control loop is closed. This results in still further improvement. With the wavefront nearly flat, white light fringes will have good contrast. The transmitted wavefront is an ideal phase reference for sensing the critical high-order ripples that would diffract starlight out as far as the planet's image.

We suppose that four phase settings at phase intervals of $\pi/2$ are used, and that imaging pixels for the two detectors are registered, each recording the light from a square subaperture of dimension Δx in the pupil. An examination of measurement errors from photon noise shows that the mean square phase error at measurement wavelength λ_s for a subaperture at radius ρ is given by

$$\sigma_p^2 = 2/(N(\rho)\eta^2(\rho)) \quad (2)$$

where ρ is the radius at the subaperture normalized to the telescope radius, $D/2$, $N(\rho)$ is the total number of photons recorded during one complete measurement from the subaperture, and $\eta(\rho)$ is the fringe visibility at that radius¹¹. The effect of removing all the Airy rings is to make the intensity of the reference beam fall off with radius. Thus σ_p^2 is highest at the edge of the

pupil, where the intensity and visibility are both weakest. From the calculated reference beam profile, I have explored the dependence of errors on division of intensity between the reference and test beam, and on broadening the test beam with a diverger. The smallest error corresponds to an even division of intensity and focal ratio of the reference beam decreased by $\sqrt{2}$. The reference beam intensity at the pupil edge for wavefront errors $\ll 1 \text{ rad}^2$ is then one-third that of the direct beam, and the intensity and visibility are given by

$$N(1) = 2/3(F_s e_s \Delta t \Delta x^2) \quad \eta(1) = \sqrt{3}/2 \quad (3)$$

where F_s is the flux from the star in photons $\text{m}^{-2} \text{s}^{-1}$ in the sensor band at λ_s , and e_s is the overall quantum efficiency, including transmission losses and detector quantum efficiency. Δt is the time allowed for the phase measurement.

The limit to wavefront measurement accuracy

Two sources of noise set the fundamental limit to wavefront measurement accuracy. The random arrival of photons results in a fractional error equal to $1/\sqrt{N}$ when N photons are detected. Clearly, to reduce photon noise, it is desirable to increase N by using larger values of Δx and Δt . However, errors are then introduced because of inadequate spatial and temporal sampling of the wavefront. Here the optimum sampling and the corresponding minimum error are derived, using the least-squares method and modelling the atmospheric wavefront aberration as Kolmogorov turbulence in refractive index.

Atmospheric turbulence introduces phase differences whose mean square value between points separated by distance x is given in radians² by

$$\sigma_x^2 = 6.88(x/r_0)^{5/3} \quad (4)$$

Here the strength of turbulence is characterized by Fried's length, r_0 , the characteristic distance beyond which phase is effectively randomized¹². From this phase structure function it may be shown that the total error including photon and atmospheric noise is given by¹³

$$\sigma^2 = 1/(F_s q \Delta t \Delta x^2) + (g \Delta x / r_0)^{5/3} + (w \Delta t / r_0)^{5/3} \quad (5)$$

Here the first term is the photon noise in terms of the dimensionless parameter q , a measure of overall photon efficiency for any sensor. For the wavefront sensor analysed above, the photon noise error depends also on location in the pupil. For the largest error which occurs at the pupil edge, it follows from equations (2) and (3) that $q = e_s/4$. To be conservative we will assume the same value of q for the whole pupil. The second term gives the strength of spatial aberration error, characterized for any sensor in terms of a second dimensionless parameter g depending on geometry. For optimum reconstruction from discrete phase measurements over a square pixel of dimension Δx we find $g = 0.4$, from the analysis of wavefront fitting by Hudgin¹⁴. The third term gives the error from temporal evolution of the wavefront appropriate for small Δx , in terms of atmospheric parameters w and r_0 . In the approximation of frozen turbulence carried by the wind, $w = 3.2v_w$, where v_w is the average wind speed weighted by strength of turbulent distortion¹⁵.

To find the ultimate accuracy, we shall suppose the wavefront sensor is configured with Δx and Δt adjusted so as to minimize the total wavefront error. By the least-squares method, setting $\partial \sigma^2 / \partial \Delta x = 0$ and $\partial \sigma^2 / \partial \Delta t = 0$, we obtain from equation (5)

$$\Delta x_{\text{opt}} = 1.14(F_s q / w)^{-3/14} g^{-4/7} r_0^{5/14} \quad (6)$$

and

$$\Delta t_{\text{opt}} = 0.75(F_s q)^{-3/14} g^{3/7} w^{-11/14} r_0^{5/14} \quad (7)$$

and with these optimum values equation (5) becomes

$$\sigma_{\text{min}}^2 = 2.9(F_s q r_0^3 / w g^2)^{-5/14} \quad (8)$$

Thus in this limit, the accuracy of measurement is characterized

by just the two dimensionless parameters g and q for the sensor, the values of w and r_0 for the atmosphere, and the star's photon flux F_s at the sensor. Substituting from equations (2) and (5), the estimate for g given above, and setting $r_0/v_w = \tau$, the timescale for temporal evolution, we obtain

$$\sigma_{\text{min}}^2 = 2.9(F_s r_0^2 \eta^2 \tau e)^{-5/14} \quad (9)$$

where e here is the quantum efficiency including light losses by diffraction at the spatial filter. This relation for wavefront error may be compared with the result of a study of photon and atmospheric noise in adaptive optics systems made by Dyson¹⁶. He found for systems of this type that the minimum photon flux for adaptive correction must satisfy the condition

$$F_s r_0^2 \eta^2 \tau e > 1 \quad (10)$$

where factors of order unity were dropped. Equation (9) thus is consistent with, and extends, Dyson's condition, giving the improvement in phase error made possible with photon fluxes higher than the minimum. Because the interferometric system analysed is representative of the most efficient wavefront sensors in terms of photon noise, equations (8) and (9) represent a practical limit to the accuracy of any adaptive optics system.

Correction for scintillation

So far we have considered only the measurement and correction of atmospheric phase disturbance. But when this is done to high accuracy, the effect of intensity fluctuations from scintillation will become dominant in stellar scattering. In this section, it will be seen that such effects, not previously been considered, must also be corrected adaptively if planets are to be detected efficiently.

The variance of the natural logarithm of the intensity fluctuations from scintillation is given by Tatarski¹⁷:

$$\sigma_{\ln I}^2 = 2.24 \left(\frac{2\pi}{\lambda} \right)^{7/6} \sec^{11/6}(z) \int C_n^2(h) h^{5/6} dh \quad (11)$$

where z is the zenith angle, h is height above the telescope, C_n^2 is the refractive index structure constant and λ is the wavelength of observation. If the phase errors were perfectly corrected, the fraction of energy scattered by amplitude errors would be equal to the normalized variance σ_a^2 of wavefront amplitude. This may be understood if we observe that scintillation errors are rotated 90° from phase errors in the Argand diagram, but otherwise sum in exactly the same way. For the small errors we encounter $\sigma_a^2 = \sigma_{\ln I}^2/4$. Supposing for simplicity that observation is made at the zenith and that the high-altitude component is localized at height h_0 and characterized in strength by $r_0(h)$, equation (11) yields

$$\sigma_a^2 = 0.29 \left(\frac{\sqrt{\lambda_m h_0}}{r_0(h)} \right)^{5/3} \quad (12)$$

The total energy scattered into a halo from uncorrected intensity fluctuations for our application would be larger than that from residual spatial phase errors (the second term in equation (5)), as the Fresnel length, the spatial scale for scintillation¹⁷, $\sqrt{(\lambda_m h_0)} > g \Delta x_{\text{opt}}$. Furthermore, as halo diameter scales inversely with spatial scale of scintillation, and this is several times larger than the phase correction scale Δx_{opt} , the larger energy from scintillation would be scattered into a smaller and thus brighter halo. We note also that the timescale for speckle change, $\sqrt{(\lambda_m h_0)}/v_w$, is typically longer than Δt_{opt} , thus averaging of the halo speckle background needed to bring out a planet (considered in the section below) would be slower than accomplished for phase errors.

For these reasons, it is important to correct intensity fluctuations. Measurements of adequate accuracy are required, and a device analogous to the adaptive phase mirror to lower intensity everywhere to the same level. Fortunately photon noise allows

the recovery of intensity fluctuations with high accuracy and resolution from the wavefront sensor. From equation (2) the phase variance for each pupil pixel measurement from photon noise is $\sim 4/N(\rho)$, whereas the amplitude variance will be $1/4N(\rho)$, much smaller. A liquid-crystal transmission device could probably be used for intensity correction.

Noise in the corrected image halo

The image of the star and planet will now be reconsidered, particularly the noise character of the background halo after correction. As in the wavefront sensor, photon noise will cause random fluctuations in intensity. So will speckle structure arising from spatial coherence of the starlight, with structure on the scale of the diffraction limit. In this section, the gain and limiting star/planet brightness ratio set by halo noise are determined.

Our supposition is that the adaptive elements used to correct wavefront phase and intensity have all the resolution and speed needed to correct the wavefront according to the sensor measurements. With no further error introduced, the residual phase error of the system at the measurement wavelength will be correctly given by equation (8). The gain at the imaging detector at wavelength λ_m is then given by equations (1), (6) and (8):

$$G = 0.27(F_s q/w)^{+11/14} g^{3/7} r_0^{5/14} D^2 (\lambda_m/\lambda_s)^2 \quad (13)$$

To determine the halo noise contribution, we must evaluate the two contributions of speckle and photon noise. The total number of stellar photons arriving at the focus in this interval is $F_m \Delta t_{\text{opt}} \pi D^2/4$, where F_m is the photon flux for imaging. Dividing by the gain, the mean number N_p of photons recorded in the halo within a single diffraction limited image pixel during the integration time Δt_{opt} time is given from equations (7) and (13) by

$$n_p = F_m e_m \Delta t \pi D^2/4G = (7\pi/5)(F_m e_m/F_s q)(\lambda_s/\lambda_m)^2 \quad (14)$$

where e_m is the effective quantum efficiency of the imager. For the case we have analysed $e_m/q \sim 4$ and $(\lambda_s/\lambda_m)^2 = 1$, thus $N_p \sim 16F_m/F_s$. It follows that on the timescale of corrections, photon noise is quite strong. For example, we will adopt $F_m = 1/4F_s$ when $N_p = 4$, and the r.m.s. fractional noise equal to $1/\sqrt{N}$ is 50%.

Speckle structure is effectively frozen between correction intervals Δt_{opt} , and has contrast of $\sim 100\%$ in monochromatic light. Averaged over a finite spectral bandwidth, the speckles are smeared in the radial direction¹⁸. At angular radius θ the r.m.s. intensity fluctuation will be approximately $\sqrt{(\theta D \Delta \lambda)/\lambda}$. For typical values given below, this fluctuation will be around 30%. To be conservative, we will assume that the combined effect of photon and speckle noise will yield an S/N ratio of unity in each correction interval. Thus a diffraction-limited planet image which is G times fainter than the star, and located within the underlying residual speckle halo, will be detected at a S/N ratio of ~ 1 in time Δt_{opt} .

To the extent that the intensity errors are uncorrelated from one correction interval to the next, the error made in halo intensity in a long-exposure image is reduced as the square root of the number of corrections, according to the central limit

TABLE 1 Properties of best candidate single stars

Star	Spectral type	Distance (pc)	m_R^* (mag)	$F_s^\dagger$ ($10^8 \text{ m}^{-2} \text{ s}^{-1}$)
α Lyr	A0V	7.5	0.0	140
α PsA	A3V	6.7	1.0	56
α Aql	A7V	4.9	0.6	86
π^3 Ori	F6V	7.3	2.8	11.5
γ Lep	F7V	7.8	3.1	8.1
ζ Tuc	F9V	7.1	3.7	4.6
β Hyi	G2IV	6.3	2.3	18
μ Her	G5IV	7.5	2.9	10.4
δ Pav	G5IV	5.7	3.0	9.3
τ Ceti	G8V	3.5	2.9	9.8
HD 20794	G8V	6.2	3.7	4.7
σ^2 Eri	K1V	4.8	3.7	4.7
ϵ Ind	K4V	3.4	3.8	4.5

* R-band apparent magnitude.

† Photon flux.

theorem. Photon noise is certainly random. Successive speckle patterns will also be uncorrelated, because each new wavefront measurement is made at thousands of independent points, and photon noise in the wavefront sensor ensures different phase errors at each point. Thus the very short correction interval works to our advantage, because a very large number of realizations of the corrected image, $T/\Delta t_{\text{opt}}$, are averaged in a long exposure. For an integration of duration T , the limiting ratio of star/planet brightness, for a detection at signal-to-noise ratio S/N is given by

$$R = G \sqrt{T/\Delta t_{\text{opt}}} / (S/N) \quad (15)$$

$$= 0.31(F_s q)^{25/28} w^{-11/28} g^{3/14} r_0^{5/28} D^2 (\lambda_m/\lambda_s)^2 T^{1/2} (S/N)^{-1}$$

Projected sensitivity for nearby stars

The relations developed above are now used to estimate the optimum system parameters and performance for bright star candidates observed with large telescopes. Good candidates for Jupiter-like planets that are both bright and nearby are given in Table 1. Extending the list of Woolf and Angel¹⁹, Table 1 lists all the stars that are within 8 pc, have apparent magnitude brighter than 4.0 in the R band, and are either single or have no stellar companion within 200 AU (ref. 20). These criteria yield ten main sequence and three slightly evolved stars (type IV), with masses not far from solar (type G2). Two of the three A stars, Vega (α Lyrae) and Fomalhaut (α Piscis Austrini) have dust discs²¹.

The wavelength band 0.65–0.95 μm ($\lambda_s = 0.8 \mu\text{m}$) is chosen for wavefront measurements, to take advantage of silicon CCDs (charge-coupled devices) with low noise and high quantum efficiency (>0.8 average in this band). Assuming that the transmission of the telescope optics is 0.8, we obtain an effective quantum efficiency $e_s = 0.65$, and thus $q = 0.16$. Because chromatic differences between even adjacent photometric bands can result in differential wavefront aberration of $\sim 0.01 \text{ rad}^2$ (ref. 22), I will assume that the imaging will be carried out in the same

TABLE 2 System parameters and performance

Star flux F_s (photons $\text{m}^{-2} \text{s}^{-1}$)	4×10^8	10^9	10^{10}
Optimum subaperture Δx_{opt} (mm)	52	43	26
Optimum measurement time Δt_{opt} (ms)	0.43	0.35	0.22
Mean square phase error σ_{min}^2 at λ_s	0.037	0.027	0.012
r.m.s. wavefront error (nm)	25	21	14
Detected photon count $N(1)$ (equation (3))	200	280	630
r.m.s. photon noise per pixel read (e)	5.0	5.9	8.9
Telescope aperture (m)	6.5	6.5	6.5
Gain $G \times 10^6$ (equation (13))	0.42	0.85	5.2
Star/planet ratio $R \times 10^6$ for $S/N=5$ in 1 h (equation (15))	240	540	4,300

waveband as the wavefront sensing. An 80%–20% beamsplitter is adopted, favouring the wavefront sensor, to approximately balance speckle and photon noise in the imaging camera. Photon fluxes F_s , allowing for the 0.3- μm bandwidth and beamsplitter loss, are listed in Table 1, and range from 4.5 to $>100 \times 10^8$ photons $\text{m}^{-2} \text{s}^{-1}$.

For atmospheric parameters and telescope diameter I take values appropriate for the telescopes with which we plan to make direct imaging planet searches, the 6.5-m Monolithic Mirror Telescope (MMT) and Magellan telescope, and the double 8.4-m Large Binocular Telescope (LBT)²³. By combining coherently the adaptively corrected images from both mirrors of the LBT, the effective collecting area and image solid angle will be the same as a single telescope with $D=11.8$ m. A value of 10 m s^{-1} is adopted for the turbulence-averaged wind speed, and r_0 is taken to be 0.25 m at 0.8 μm , as recently measured at the MMT (M. Lloyd-Hart *et al.*, manuscript in preparation).

The parameters given above are sufficient to find the optimum design and resultant photon count and phase error. These are given in Table 2 for different fluxes F_s appropriate for stars in Table 1. Also listed for both telescope apertures is the gain from equation (13) and the star/planet intensity ratio predicted from equation (15) to give an S/N ratio of 5 s.d. in a 1-h integration. Note that practical realization of the listed gain requires uncorrected wavefront aberrations on spatial scales less than Δx_{opt} , around 5 cm, to be less than the projected ~ 20 nm r.m.s. correction error. Fortunately this is the case for the best current ground-based optics²⁴.

The projections of Table 2 are very encouraging. When seen from afar, half-illuminated at maximum extension, a Jupiter-like planet with 5-AU orbit would be $\sim 8 \times 10^8$ times fainter than its host star⁴. The LBT has the sensitivity to reach this contrast ratio for any of the candidates in a 1-h integration, at the 5 s.d. level. In 2–3 h the smaller 6.5 m telescopes will reach the same performance for most of the stars in Table 1. The faintest ($F_s \approx 4 \times 10^8$) will require 10 h, given the dependence of the star/planet brightness ratio on the square root of the integration time. Thus detection of Jupiter-like planets at 5 s.d. should be possible for all the listed stars, given the system parameters listed in Table 2 optimized for lowest fluxes, that is, sampling at 2,000 Hz with 6-cm pixels.

The maximum angular separation of a planet with the same size and orbit as Jupiter is 1.5 arcsec at the ~ 3.4 pc distance of the closest candidates, τ Ceti and ϵ Indi, and 0.65 arcsec for the most distant star in Table 1. Because halo background is not centrally peaked, angular separation does not enter into our calculation of limiting sensitivity. But care must be taken at the

smaller separations to reduce the effects of aperture diffraction by apodization, that is shading at the pupil edges to remove sharp intensity transitions.

A 6.5-m telescope corrected with 3-cm sampling at 4-kHz rate, the optimum listed in Table 2 for high fluxes of $\geq 10^{10}$ photons $\text{m}^{-2} \text{s}^{-1}$, should reach brightness ratios of $\sim 4 \times 10^9$ in 1 h. Small planets in stable close orbits could thus be detected for the listed bright A stars and also for the three very bright close binaries with separations of ~ 10 AU, α Centauri (1.3 pc), Sirius (2.8 pc) and Procyon (3.4 pc)²⁵. The sensitivity of the 6.5-m Magellan telescope will be enough to see in a few hours an Earth-sized planet of α Cen A (type G2) at 1 AU radius (0.75 arcsec separation, 6×10^9 contrast ratio²⁶).

Given the need for observations of each star with different telescope orientations to average systematic errors, to be continued over years to find favourable maximum separations, a survey of the dozen candidates given above would be about the practical limit with general use telescopes. With a telescope dedicated to the task and integrations 10 times longer, from equation (14) we project stars 3.6 times fainter could be reached at given brightness ratio. This corresponds to a factor 1.9 increase in distance, or a sevenfold increase in volume, yielding ~ 100 candidates. However, observations with lower gain of fainter objects at smaller angular separation may tax to the limit our ability to control systematic errors.

In conclusion, the main finding of this study is that photon noise will be small enough to allow a good sampling of nearby stars for extrasolar planets with the largest telescopes now under construction. This would not be possible because of measurement uncertainty if the photon fluxes were only a hundred times smaller. The technology for correction with high accuracy and speed developed by the US military over the past 20 years goes a long way toward meeting the requirements for a practical system²⁷. In a more extended discussion (J.R.P.A., manuscript in preparation), I consider how systematic errors may be held below the very low random errors projected by this work. The ground-based search of the nearest stars is visualized as a precursor to a powerful space-based programme, with complementary technical challenges. In space, errors are quasi-static, but their calibration and the suppression of diffraction by small apertures will probably be problematic⁴. Adaptively corrected ground-based telescopes avoid these difficulties because of their large aperture and rapid sensing and randomizing of residual errors. It seems likely that by the end of the decade ground-based observations will either have found planets around nearby main-sequence stars, or have shown conclusively that planets of the size of Jupiter are rare. $\square$

Received 1 November 1993; accepted 14 February 1994.

- Wolszczan, A. & Frail, D. A. *Nature* **355**, 145–147 (1992).
- Angel, J. R. P., Cheng, A. Y. S. & Woolf, N. J. *Nature* **322**, 341–343 (1986).
- Angel, J. R. P. in *Proc. Workshop on The Next Generation Space Telescope* (eds Bely, P. & Burrows, C. J.) 81–94 (Space Telescope Science Institute, Baltimore, 1990).
- Brown, R. A. & Burrows, C. J. *Icarus* **87**, 484–497 (1990).
- Beckers, J. A. *Rev. Astr. Astrophys.* **31**, 13–62 (1993).
- Born, M. & Wolf, E. *Principles of Optics* (Pergamon, Oxford, 1980).
- Angel, J. R. P. in *Proc. Laser Guide Star Adaptive Optics Workshop Vol. 2* (ed. Fugate, R. Q.) 494–499 (Phillips Lab, Albuquerque, 1992).
- Creath, K. *Progress in Optics Vol 26* (ed. Wolf, E.) 349–393 (Elsevier, New York, 1988).
- Zernike, F. *Mon. Not. R. astr. Soc.* **94**, 377–384 (1934).
- Colucci, D. Thesis, Univ. Arizona (1993).
- Wyant, J. *Appl. Opt.* **14**, 2622–2626 (1975).
- Fried, D. L. *J. opt. Soc. Am.* **56**, 1373–1379 (1966).
- Angel, J. R. P. in *Proc. Institut D'Etudes Scientifiques de Cargèse* (eds Alloin, D. & Mariotti, J.-M.) (NATO Series C, Kluwer, The Netherlands, in the press).
- Hudgin, R. J. *J. opt. Soc. Am.* **67**, 393–395 (1977).
- Sandler, D. G., Stahl, S., Angel, J. R. P., Lloyd-Hart, M. & McCarthy, D. J. *J. opt. Soc. Am.* (in the press).
- Dyson, F. J. *J. opt. Soc. Am.* **65**, 551–558 (1975).
- Tatarski, V. I. *Wave Propagation in a Turbulent Medium* (McGraw-Hill, New York, 1961).

- Dainty, J. C. (ed.) *Laser Speckle and Related Phenomena* 272–274 (Topics in Applied Physics Vol. 9, Springer, Berlin, 1984).
- Woolf, N. & Angel, J. R. P. in *Astrophysics from the Moon* (eds Mumma, M. J. & Smith, H. J.) 95–98 (AIP Conf. Proc. Vol. 207, Am. Inst. Phys., New York, 1990).
- Hoffleit, D. *Yale Catalog of Bright Stars* 3rd edn (New Haven Printing, New Haven, Connecticut, 1964).
- Backman, D. E. & Paresce, F. *Protostars and Planets III* (eds Levy, E. H. & Lunine, J. E.) 1253–1304 (Univ. Arizona Press, Tucson, 1993).
- Hogge, C. R. & Butts, R. R. *J. opt. Soc. Am.* **72**, 606–609 (1982).
- Martin, H. M., Hill, J. M. & Angel, J. R. P. *Physics Today* **44**, March, 22–30 (1991).
- Anderson, D. *et al.* in *Proc. Soc. Photo-opt. Instrumental Engrs* (ed. Doherty, V. J.) 193–207 (SPIE, Bellingham, Washington, 1994).
- Aitken, R. G. *The Binary Stars* (Dover, New York, 1964).
- Burke, B. F. *Nature* **322**, 340–341 (1986).
- Fugate, R. Q. *et al.* in *Proc. Laser Guide Star Adaptive Optics Workshop Vol. 1* (ed. Fugate, R. Q.) 132–163 (Phillips Lab, Albuquerque, 1992).

ACKNOWLEDGEMENTS. I am grateful to N. Woolf for alerting me to the scattering caused by scintillation, and also to M. Lloyd-Hart, D. Sandler and J. Burge for discussions on adaptive optics, and their readings of the manuscript. Adaptive optics research at Steward Observatory is supported by the US NSF and the Flintridge Foundation.

Compartment boundaries and the control of *Drosophila* limb pattern by hedgehog protein

Konrad Basler* & Gary Struhl†

* Zoologisches Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

† Howard Hughes Medical Institute, Columbia University College of Physicians and Surgeons, 701 West 168th Street, New York, New York 10032, USA

***Drosophila* limbs are subdivided into anterior and posterior compartments which derive from adjacent cell populations founded early in development. Evidence is now provided that posterior cells organize growth and cell patterning in both compartments by secreting hedgehog protein and that hedgehog protein acts indirectly by inducing neighbouring anterior cells to secrete decapentaplegic or wingless protein.**

EACH adult appendage of *Drosophila* is subdivided into anterior and posterior compartments, precisely defined portions of the integument formed by the descendants of groups of founder cells established early in development¹⁻³. Compartmentalization involves the heritable activation of the *engrailed* (*en*) gene in all of the posterior founder cells⁴⁻¹⁰. Once active, *en* functions continuously to prevent intermixing between posterior and anterior cells, and to direct posterior cells to form posterior rather than anterior patterns^{4,5,11}.

Although much attention has centred on the apparent autonomy of anterior and posterior compartments, it has long been appreciated that the boundaries between these compartments might have special organizing properties^{2,12}. In particular, interactions between anterior and posterior cells at the boundary may lead to the local synthesis¹² or directed transport² of signalling molecules which organize cellular behaviour as a function of distance from the boundary.

Three genes, *hedgehog* (*hh*)¹³, *decapentaplegic* (*dpp*)¹⁴ and *wingless* (*wg*)¹⁵, which encode secreted proteins¹⁶⁻²¹, may be involved in this signalling process. All three have been implicated as intercellular signals in *Drosophila*^{18,22-34} and belong to conserved gene families encoding growth or differentiation factors in other animals³⁵⁻⁴⁰. *hh* is expressed coincidentally with *en* in all posterior compartment cells^{20,41-43}. In contrast, *dpp*^{27,44-46} and *wg*^{34,47} are expressed in anterior cells that are neighbours to the *hh*-expressing posterior cells at the compartment boundary. Thus, cells on either side of the compartment boundary could exert an organizing influence on surrounding tissue by providing local sources of one or more of these putative signalling molecules.

Here we show that ectopic expression of *hh* can induce ectopic *dpp* or *wg* expression in anterior cells and reorganize anterior compartment pattern. Conversely, loss of endogenous *hh* function blocks *dpp* and *wg* expression along the compartment boundary and impedes growth and patterning in both compartments. We propose that (1) posterior cells organize limb patterning by secreting *hh* protein, and (2) *hh* acts indirectly by inducing *dpp* or *wg* expression in anterior cells that abut the compartment boundary.

hh organizes anterior compartment pattern

We have used the 'flp-out' technique³⁴ to activate constitutive *hh* expression in genetically marked clones of cells positioned at random in the imaginal discs. As described in Fig. 1, this technique uses the yeast site-specific recombinase *flp*⁴⁸ to fuse the constitutive *Tubulinal* (*Tuba1*) promoter⁴⁹ to the *hh* coding sequence by excising a segment of DNA containing the *yellow*⁺

(*y*⁺) marker gene. Animals mutant for the endogenous *y* gene but carrying a *Tuba1* > *y*⁺ > *hh* transgene ('>' indicates targets for *flp*-mediated recombination) as well as an *hsp70-flp* transgene were heat-shocked during early larval development. Clones of *y*; *Tuba1* > *hh* cells were then identified in adults by the presence of cuticular structures such as bristles and hairs which were phenotypically yellow.

In the wing, the anterior and posterior compartments form precisely defined sub-patterns of hairs, veins and sense organs (such as bristles) separated by a boundary positioned just anterior to vein 4 (ref. 1) (Fig. 2a). Both compartments are further subdivided into dorsal and ventral compartments which form the dorsal and ventral surfaces of the wing blade¹. Clones of *y*; *Tuba1* > *hh* cells were identified by the presence of *y* bristles

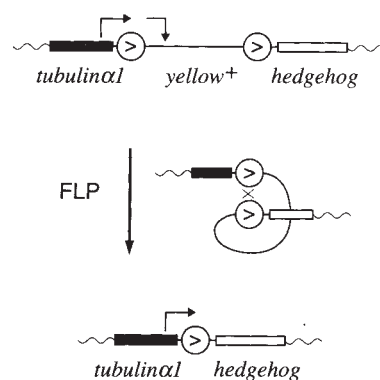


FIG. 1 Constitutive activation of *hh* expression by *flp*-mediated recombination within the *Tub* > *y*⁺ > *hh* transgene. The *Tub* > *y*⁺ > *hh* transgene consists of a 2.6 kilobase (kb) fragment from the promoter of the *Tuba1* gene⁴⁹, which drives constitutive expression in most cells throughout development (K.B. and G.S., unpublished data), positioned upstream of the *hh* coding sequence (position 221 to 2,109)¹², but separated from this coding sequence by a > *y*⁺ > *flp*-out cassette which blocks transcription. The *flp*-out cassette contains the intact *y*⁺ gene and is flanked by targets for the yeast site specific recombinase *flp*³⁴. These *flp* recombination targets (termed FRTs and indicated by '>' in the text) have the same orientation. As shown, exposure to the *flp* recombinase can cause a recombination event which leads to excision of the intervening DNA leaving behind only a single FRT. The transgene is inserted in the C20NX (P[ry⁺]) transformation vector³⁴. Further details about the construction of this transgene are available on request.

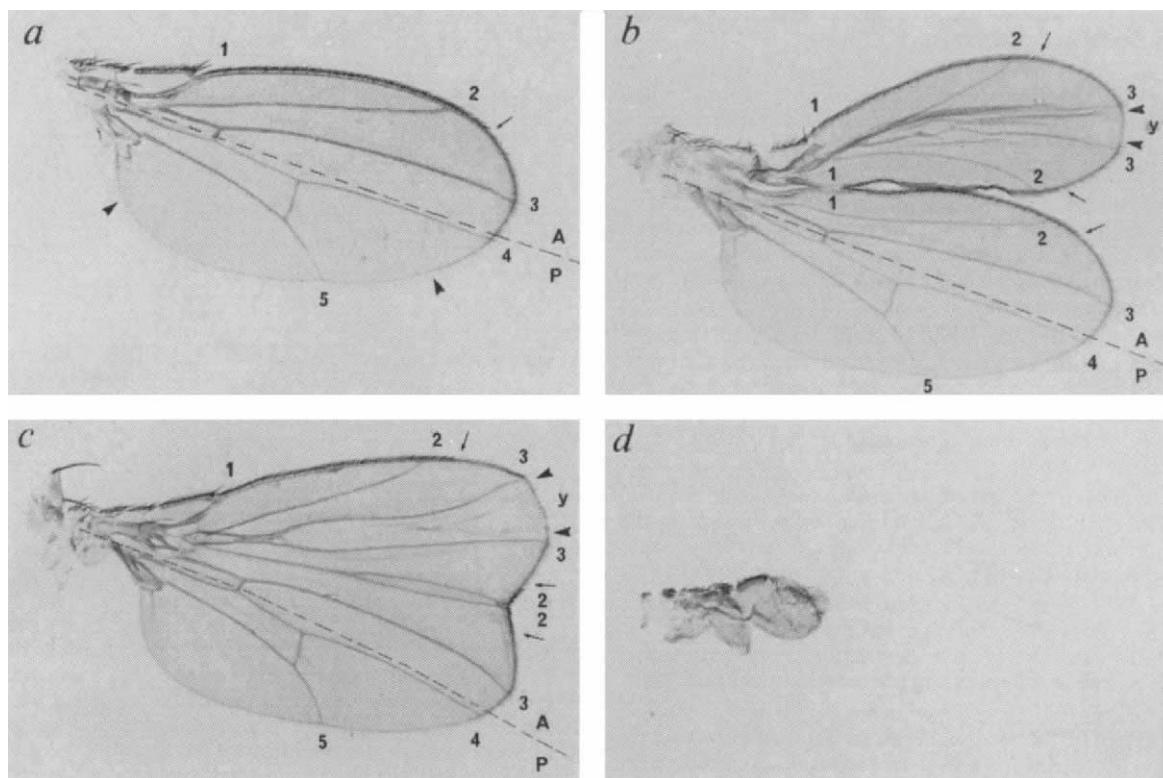


FIG. 2 Reorganization of anterior wing pattern by genetically marked clones of cells that express *hh* constitutively. *a*, Normal wing pattern. A dotted line marks the position of the anterior-posterior compartment boundary which runs just anterior to vein 4 (longitudinal veins 1–5 are numbered at their point of intersection with the wing margin). The anterior portion of the wing margin displays a distinct pattern of three rows of bristles (termed the 'triple row') which extends along the margin to a position just posterior to the intersection with vein 2; the remainder of the wing margin carries a 'double row' of long, slender hairs (a small arrow marks the position of the transition from triple to double row). This wing carries a large clone of *y; Tuba1>hh* cells in the posterior compartment which contributes most of the hairs along the posterior wing margin (arrow heads mark the boundaries of *y* tissue along the wing margin). *b*, *c*, Alterations of anterior compartment pattern associated with *y; Tuba1>hh* clones in the anterior compartment. In each case, arrow heads mark the boundaries of *y; Tuba1>hh* tissue along the wing margin; the longitudinal veins, triple row/double row transition, and normal anterior-posterior compartment boundary are indicated as in *a*. Note that a nearly complete anterior compartment pattern always forms anterior to the clone and that the remaining portion of the anterior compartment contains symmetrical anterior compartment patterns bounded by the clone and the normal compartment boundary. These two cases differ in the extent of the mirror symmetric anterior pattern formed between the clone and the compartment boundary, possibly due to differences in the position of the founding cell of the clone relative to the compartment boundary. Note that the more extensive symmetric pattern shown in *b* also correlates with the formation of a supernumerary wing composed of two anterior compartment patterns organized in mirror symmetry relative to the clone. *d*, Reduction of both anterior and posterior compartments associated with a *hh* mutant clone in the posterior wing compartment. This clone marks bristles on the posterior portion of the wing blade, and the wing consists of only extreme posterior and extreme anterior structures. Comparing *b* and *d* it is apparent that gain or loss of *hh* function is associated not only with gain or loss of anterior compartment pattern, but also with gain or loss of proximal-distal pattern. We suggest that the effects on the proximal-distal axis are mediated indirectly by the effects of altered *hh* activity on anterior-posterior patterning.

METHODS. Generation of *y; Tuba1>hh* clones: *y hsp70-flp* females were crossed to males carrying a single insert of the *Tuba1>y⁺>hh* gene on chromosome II, the resulting larvae subjected to a mild heat shock (33, 33.5 or 34 °C for 30 or 60 min) during the first or second larval instar (24–48 or 48–72 h after fertilization) and

y hsp70-flp/Y; Tuba1>y⁺>hh/+ adults collected and screened for *y; Tuba1>hh* clones. Most clones in the anterior compartment were easily recognized because they were associated with altered venation patterns. Posterior compartment clones and phenotypically normal anterior compartment clones were detected by screening the remaining, normally patterned wings under the compound microscope for *y* bristles or hairs on the wing margin. For example, in a sample of 68 wings obtained from flies derived from first instar larvae subjected to a 30 min heat shock at 34 °C, 15/15 clones marking the wing margin in the posterior compartment were associated with normal patterning, whereas 7/13 clones marking the margin in the anterior compartment were associated with abnormal venation patterns (of the 6 anterior compartment clones which appeared normal, 4/6 marked margin hairs close to vein 4, a region next to the compartment boundary and hence close to *hh*-expressing posterior cells; the remaining 2/6 marked only one or a few margin hairs, and hence may have been small, spontaneous clones which arose late in development). An additional 9 cases of altered venation were observed in the anterior compartments of these wings, although they were not associated with *y* tissue at the wing margin (no cases of abnormal venation were observed in the posterior compartments of the 68 wings in this sample). Variations in heat-shock conditions and the timing of heat shock influenced the number and size of clones as observed previously³⁴. Wings were dissected in propanol and mounted in a 1:1 mixture of Canada balsam and methyl salicylate⁶⁵; all are shown at the same magnification. Generation of clones lacking *hh* gene function. Two techniques, *flp*-mediated mitotic recombination^{66,67} and the *Minute* technique⁶⁸ were used together to generate wings carrying large clones of cells lacking *hh* gene function. *hsp70-flp; FRT-82(neoR) Sb^{63b}M(3)w¹²⁴/TM2* females were crossed to *FRT-82(neoR) hh^{E23} e¹¹/TM6B* (experimental) or *FRT-82(neoR) e¹¹* (control) males, the resulting larvae subjected to a moderate heat shock (34 °C for 60 min) during the first larval instar, and *FRT-82(neoR) hh^{E23} e¹¹/FRT-82(neoR) Sb^{63b}M(3)w¹²⁴* (experimental) or *FRT-82(neoR) e¹¹/FRT-82(neoR) Sb^{63b}M(3)w¹²⁴* (control) progeny (all of which carry a single copy of the *hsp70-flp* gene) screened for altered wing patterns. Of 180 experimental progeny obtained, 16 showed a severe reduction in one wing, and in 7/16 cases this reduction was associated with phenotypically *Sb⁺ e¹¹ M(3)w⁺* bristles on the remaining portion of the posterior wing margin. In contrast, 0/350 control progeny showed severely reduced wings. Similar results were also obtained with a second *hh* allele, *hh^{AC}* (the *hh^{E23}* and *hh^{AC}* mutations were induced on different parental chromosomes in separate genetic screens^{20,33}).

or hairs at the wing margin which coincides with the boundary between the dorsal and ventral compartments.

Two classes of $y; Tuba1 > hh$ wing clones were observed (see Fig. 2 legend for quantification). Clones belonging to the first class fell within the posterior compartment and, aside from the y phenotype, contributed to the formation of a normal posterior compartment (Fig. 2a). All clones in the second class fell within the anterior compartment and were usually associated with abnormal vein patterns (Fig. 2b, c). Although the abnormal patterns observed in the wing varied depending on the position of the clone, they can be described by two simple rules: (1) cells belonging to the clone develop like normal anterior cells adjacent to the compartment boundary and organize wild-type cells situated anteriorly to form an almost complete anterior pattern; and (2) wild-type cells situated posterior to the clone form symmetric anterior patterns.

Accordingly, if the normal wing pattern is described as 123/45 (in which the numbers define veins and '/' defines the compartment boundary), the abnormal pattern shown in Fig. 2b could be described as 123y321123/45 (in which 'y' marks the position of the clone), whereas less extreme pattern abnormalities could be described as 123y3223/45 (Fig. 2c) and 123y33/45 (not shown). Because ectopic expression of hh by anterior cells can organize a symmetric 123y321 pattern, we infer that the normal anterior pattern (123/) is organized by hh protein expressed by posterior cells abutting the compartment boundary. Likewise, the symmetry of the y321123/, y3223/ and y33/ patterns generated between the clone and the compartment boundary suggests that $y; Tub > hh$ cells in the clone exert an equal influence of opposite polarity to that of wild-type hh -expressing cells in the posterior compartment.

Similar results were also obtained in the legs. $y; Tuba1 > hh$ clones in the posterior compartment appear normal. In contrast, the induction of such clones in the anterior compartment caused reiterations of either dorsal anterior or ventral anterior patterns, or the formation of supernumerary legs composed of symmetric anterior compartments (Fig. 3). In the latter case, $y; Tuba1 > hh$ cells were usually found in both dorsal and ventral portions of the supernumerary limb, suggesting that supernumerary limbs arise only when hh is expressed ectopically by both dorsally and ventrally situated cells.

hh induces expression of dpp or wg protein

In principle, hh protein secreted by posterior cells might organize anterior pattern directly by diffusing in a graded fashion across the anterior compartment and specifying different outcomes at different concentration thresholds. Alternatively, it might function indirectly by inducing neighbouring anterior compartment cells to synthesize other signalling molecules which themselves act as gradient morphogens, or by initiating a longer chain of inductive events.

A thin stripe of anterior cells abutting hh -expressing posterior cells at the compartment boundary express dpp in the wing disc (Fig. 4a), and dpp or wg in the leg discs (Fig. 4c, e)^{27,34,44–47}. Hence, hh -expressing cells might normally organize anterior pattern by inducing dpp or wg expression in neighbouring anterior cells. To test whether hh protein is sufficient to induce anterior cells to express these other genes, we subjected larvae carrying the $Tuba1 > y^+ > hh$ and $hsp70-flp$ transgenes as well as $dpp-lacZ^{45}$ or $wg-lacZ^{50}$ reporter genes to a severe heat shock sufficient to join the $Tuba1$ promoter to the hh coding sequence in most cells (see legend to Fig. 4).

Spatially indiscriminate expression of hh under the control of the $Tuba1$ promoter induces expression of the $dpp-lacZ$ reporter gene throughout the anterior compartment of the wing disc, without inducing $dpp-lacZ$ expression in the posterior compartment (Fig. 4b). Hence, all anterior wing cells can be induced to express dpp in response to hh protein, whereas posterior cells are refractory.

In the leg discs, dorsal anterior cells abutting the compartment

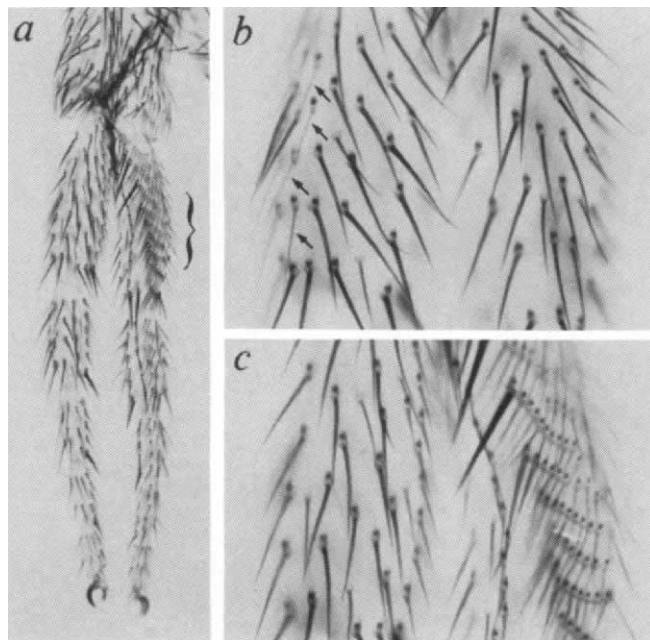


FIG. 3 Reorganization of anterior leg pattern by ectopic hh -expressing cells. Supernumerary third leg formed in association with $y; Tuba1 > hh$ clones on both the dorsal and ventral surface. The normal (right) and supernumerary (left) appendage are shown at low magnification in a. The basitarsus of the normal third leg (bracket) forms a characteristic pattern of transverse bristle rows in the posterior compartment not present in the anterior compartment. Both appendages are complete in that they consist of the normal series of tarsal segments and bear two claws and pulvilli at the distal end (not shown). b, c, Anterior and posterior surfaces, respectively, of the basitarsi of the supernumerary and normal legs at higher magnification. Note that the supernumerary appendage does not form transverse bristle rows on either the anterior or posterior surface, indicating that it is a double-anterior appendage. Arrows in b indicate $y; Tuba1 > hh$ bristles positioned dorsally on the supernumerary appendage. $y; Tuba1 > hh$ bristles are also present ventrally just proximal to the basitarsus (not shown). Clones were generated as described in Fig. 2.

boundary express dpp , whereas ventral anterior cells next to the compartment boundary express wg , as well as low levels of dpp (Fig. 4c, e). As in the wing, all anterior cells in the leg appear responsive to hh protein. But dorsal and ventral cells respond differently, dorsal cells by expressing high levels of dpp alone, and ventral cells by expressing wg as well as low levels of dpp (Fig. 4d, f).

Thus, all anterior cells have the capacity to respond to hh protein by expressing high levels of dpp or wg . But the nature of the response depends on disc type and relative position within the disc.

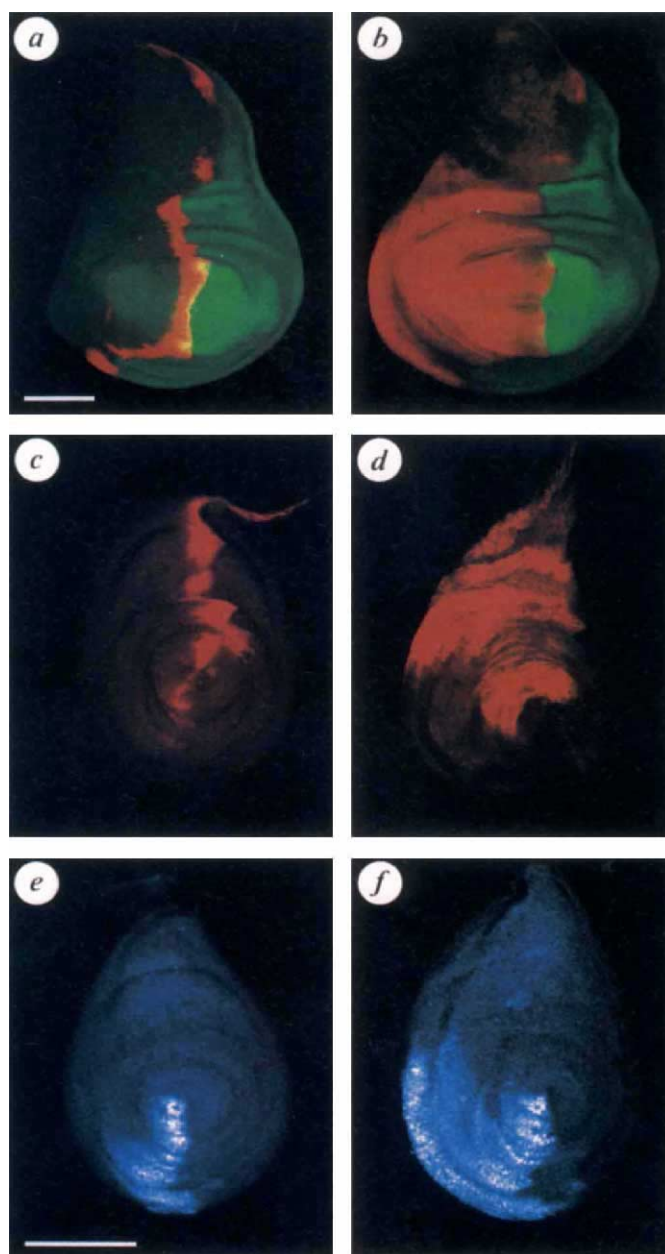
Anterior and posterior compartment development

In addition to its effects on dpp and wg expression and on anterior compartment pattern, ectopic hh activity also causes an abnormal expansion of the population of anterior compartment cells in both the wing and leg discs, without affecting the posterior compartment (Fig. 5a, b). Thus, hh -expressing posterior cells may normally stimulate cell proliferation in the anterior compartment, possibly by inducing dpp or wg protein expression. In principle, dpp or wg protein secreted by anterior cells next to the compartment boundary could also stimulate growth and organize patterning in the posterior compartment.

We have tested these possibilities by examining the development of wing and leg discs lacking endogenous hh gene function

FIG. 4 Consequences of spatially indiscriminate *hh* expression in the wing and leg imaginal discs. **a, c, e**, Control (non-heat shocked) *Tuba1>y⁺>hh* discs. **b, d, f**, Experimental (heat-shocked) *Tuba1>y⁺>hh* discs. *en* expression is shown in green, *dpp-lacZ* expression is shown in red, and *wg-lacZ* expression is shown in blue. All discs are shown dorsal side up and anterior side to the left. The leg discs (**c–f**) are shown at 1.5× the magnification of the wing discs in **a, b**; scale bars, 100 μm. In the normal wing disc (**a**), the *dpp-lacZ* reporter gene is expressed in a thin stripe of anterior cells which abuts *en*-expressing cells along the compartment boundary. In contrast, spatially indiscriminate expression of *hh* under the control of *Tuba1* causes *dpp-lacZ* to be expressed in most or all anterior cells (**b**). In normal leg discs (**c, e**), *dpp-lacZ* is expressed in a thin stripe of anterior cells which runs along the compartment boundary; however, high levels of expression are observed only in the dorsal half of the disc (**c**). Conversely, *wg-lacZ* is expressed in the ventral half of the disc, also in a stripe of anterior cells along the compartment boundary (**e**). In leg discs that express *hh* indiscriminately (**d, f**), the *dpp-lacZ* gene is expressed at high levels in most or all dorsal anterior cells, and at low levels in most or all anterior ventral cells (**d**), whereas *wg-lacZ* is expressed in most or all ventral anterior cells (**f**). Note that in the experimental discs the patterns of *dpp-lacZ* (**d**) and *wg-lacZ* (**f**) expression appear to respect a common dorsal–ventral boundary.

METHODS. Imaginal discs from control (non-heat shocked) and experimental (heat-shocked) larvae carrying *hsp70-flp³⁴*, *tub>y⁺>hh* and either *dpp>lacZ⁴⁵* or *wg-lacZ⁵⁰* constructs were fixed, immunostained for *lacZ* protein, and in the case of the wing discs for *en* protein, and examined by confocal microscopy as described previously (rabbit α -*lacZ* (Cappel) and mouse α -*en* (mAb4D9)⁶⁹ were used in conjunction with Texas red α -rabbit and FITC α -mouse secondary antisera). All discs were derived from larvae 24 h after a severe heat shock (38 °C for 60 min) which should fuse the *Tuba1* promoter to the *hh* coding sequence in most cells. *en* staining, as well as disc landmarks such as the stalk, were used to orient the discs; however, *en* staining in the legs was too weak to image photographically, and hence, is not shown.



due to the temperature-sensitive *hh^{ts2}* mutation³³ or to the generation of somatic clones of cells homozygous for apparent null mutations of *hh*^{30,33}. We find that imaginal discs of *hh^{ts2}* larvae grown at the non-permissive temperature during the third instar are abnormally small when compared to discs of *hh^{ts2}/+* siblings shifted to the non-permissive temperature at the same time. Moreover, the block in growth affects both the anterior and posterior compartment similarly (Fig. 5e, f), and correlates with the loss of *dpp* and *wg* expression, respectively, by anterior wing and leg disc cells at the compartment boundary (Fig. 5d, g).

Loss of *hh* function also appears to block normal patterning. Clones of *hh* mutant cells generated in early wing discs by mitotic recombination can block the development of large portions of both the anterior and posterior wing pattern (Fig. 2d). This phenotype, which correlates with the loss of *dpp* expression in anterior cells along the compartment boundary of *hh^{ts2}* wing discs (Fig. 5g), closely resembles that caused by eliminating *dpp* function in these same cells^{27,44}. Hence, in the wing, *hh* may normally organize growth and patterning in both compartments by inducing *dpp* expression in anterior cells adjacent to the compartment boundary. Similarly in the legs, *hh* protein may control the development of both compartments by inducing dorsally situated anterior cells to express *dpp* and ventrally situated anterior cells to express *wg*.

Compartments and limb patterning

The subdivision of the early *Drosophila* embryo into segmental units precedes the establishment of distinct populations of anterior and posterior founder cells, a process involving the heritable activation of the selector gene *en* in the posterior founder cells^{9,10,23,51–53}. Thereafter, *en* functions continuously, differentiating posterior cells from anterior cells^{4,5,11}. Based on our results, we propose that the juxtaposition of such epigenetically distinct cells along the compartment boundary provides the conditions

necessary to organize limb development. Below we outline a mechanism by which this organization may be achieved.

As illustrated in Fig. 6, we posit that posterior cells are programmed to secrete *hh* protein, in contrast to anterior cells, which are programmed to respond to *hh* protein by secreting other signalling molecules, particularly *dpp* or *wg* protein, depending on disc type and position within the disc. As a consequence, posterior cells induce a thin stripe of anterior cells along the compartment boundary to provide a source of *dpp* or *wg* protein and these proteins then act on neighbouring cells on both sides of the stripe to organize growth and patterning in each compartment. In essence, we are proposing that *hh* protein organizes limb pattern by acting as an inducer of *dpp* or *wg* expression, rather than as a gradient morphogen which specifies a series of distinct outcomes at different concentrations.

The operation of this mechanism is clearest in the wing disc because *hh* appears to act exclusively in this disc through its ability to induce *dpp* expression. We have established the organizing role of *hh* protein by showing that ectopic expression of *hh* protein by anterior cells can organize symmetrical duplica-

tions of anterior pattern (Fig. 2), whereas the loss of endogenous *hh* function blocks proliferation and patterning in both compartments (Figs 2 and 5). Similarly, we have demonstrated that all anterior cells can respond to spatially indiscriminate *hh* expression by expressing high levels of *dpp* (Fig. 4), in contrast to normal discs in which the expression of *dpp* at the boundary depends on endogenous *hh* function (Fig. 5). Thus, the secretion of *hh* protein by posterior cells appears to induce anterior cells at the compartment boundary to express *dpp* and to organize growth and patterning in both compartments.

The correlation between the organizing activity of *hh* protein and its ability to induce *dpp* suggests that *hh* controls wing patterning by directing *dpp* expression along the compartment boundary. This possibility is further supported by the finding that loss of *dpp* gene function in anterior cells adjacent to the compartment boundary causes a similar phenotype to that caused by loss of *hh* function, namely a block in normal proliferation and patterning in both wing compartments²⁷. If we assume

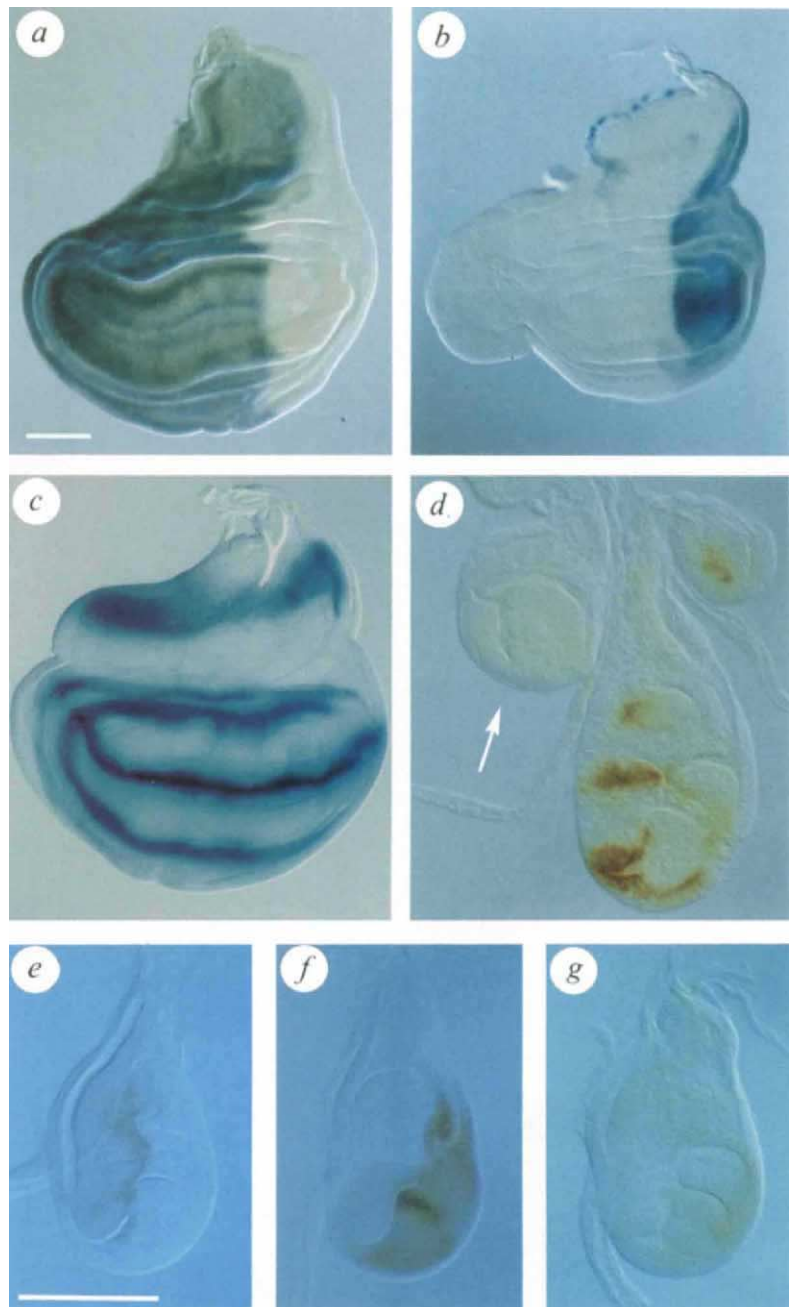
that *hh* expression by posterior cells is not adversely affected by the loss of *dpp* function in anterior cells next to the boundary, then this finding provides a strong argument that *hh* protein controls wing development indirectly through its ability to induce *dpp* expression.

The mechanism by which *hh*-expressing posterior cells organize leg pattern is more complex because anterior cells along the compartment boundary express *dpp* at high level only dorsally, whereas ventrally they express *wg*. Nevertheless, a unifying explanation emerges if we assume that *dpp* and *wg* have equivalent roles, respectively, in organizing dorsal and ventral leg pattern. Thus, we argue that the organizing influence of *hh* protein, as demonstrated by the ability of ectopic *hh*-expressing cells to direct the development of double-anterior supernumerary legs (Fig. 3), results from the induction of *dpp* and *wg* in dorsal and ventral portions, respectively, of the anterior compartment.

The situation in the leg disc is particularly informative, because we have previously shown that ectopic *wg* activity in the

FIG. 5 Involvement of *hh* protein in growth control and the localized expression of *dpp* and *wg* along the compartment boundary. *a-c*, Imaginal wing discs that express *hh* ectopically in the anterior compartment stained for *ciD-lacZ* (*a*), *hh-lacZ* (*b*) and *wg-lacZ* (*c*) expression, respectively; *d-g*, *hh^{ts2}* leg and wing discs grown at non-permissive temperature and immunostained for *wg* (*d*), *ci^D* (*e*), *en* (*f*) and *dpp* (*g*) protein expression. *a-c* are shown at the same magnification in contrast to *d-g* which are shown at twice this magnification (scale bars, 100 μ m). *a, b*, Ectopic *hh* expression induced early in development causes excessive proliferation of cells in the anterior but not in the posterior compartment of the wing disc (expression of the *ci^D-lacZ* and *hh-lacZ* genes in these discs serves as a means to visualize the anterior and posterior compartment cells, respectively^{20,70,71}). Note that *hh-lacZ* expression, which results from an enhancer trap insertion in the endogenous gene, is not altered by ectopic *hh* expression. *c*, Ectopic *hh* expression induced early in wing disc development does not noticeably alter the pattern of *wg-lacZ* expression (wild-type pattern not shown, see refs 47, 55). *d*, No *wg* expression is observed in the leg disc (left, marked by arrow) in the absence of endogenous *hh* function; however *wg* staining is observed in the wing disc (centre) as well as the haltere disc (right). *e-g*, The anterior and posterior compartments of wing discs lacking endogenous *hh* function are much smaller than in wild-type discs (*ci^D* and *en* protein staining mark the anterior and posterior compartments, respectively); in addition, *dpp* is not expressed along the compartment boundary (*g*).

METHODS. For the ectopic expression experiments, males homozygous for *hsp70-flp* and *Tuba1>y⁺>hh* transgenes were crossed to females homo- or heterozygous for the various enhancer-detector chromosomes (*ci^D-lacZ¹*; *hh-lacZ* P30²⁰; *wg-lacZ⁵⁰*), the resulting first instar larvae subjected to a severe heat shock (38 °C for 30 min), and the wing discs obtained at the end of the third instar. For the *hh^{ts2}* experiments, we fixed and immunostained imaginal discs from *hh^{ts2}* late third instar larvae selected from their *hh^{ts2}/TM6B* siblings as described previously³³. These had been shifted from 18 °C to 28 °C 60 h before fixing and staining. *lacZ* activity staining (*a-c*) was as described previously⁴⁵. Immunostaining (*d-g*) was as described previously³⁴ using rabbit α -*wg*, rabbit α -*dpp*¹⁷, rat α -*ci^D* (a gift from R. Holmgren), and mouse α -*en* (mAb4D9)⁶⁹, and HRP-conjugated secondary antisera with the appropriate host specificity.



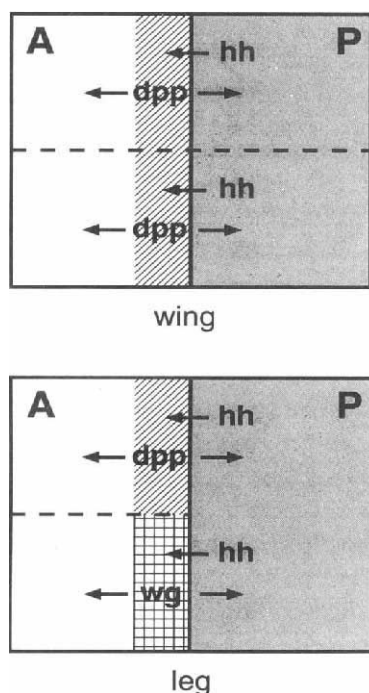


FIG. 6 Compartments and the organization of wing and leg development by *hh* protein. Both the wing and leg imaginal discs are subdivided into anterior (unshaded) and posterior (shaded) compartments. All posterior cells but no anterior cells express *en*. *en* activity programmes posterior cells to secrete *hh* protein, whereas the absence of *en* function in anterior cells programmes them to respond to *hh* protein by secreting *dpp* or *wg* protein. As a consequence, posterior cells induce a thin stripe of neighbouring anterior cells along the compartment boundary to secrete high levels of *dpp* in the wing disc and dorsal half of the leg disc (hatched) and *wg* in the ventral half of the leg disc (gridded). These anterior cells then serve as a local source for *dpp* or *wg* proteins which in turn govern growth and patterning of neighbouring cells positioned anterior and posterior to the stripe. Because cells in the anterior and posterior compartment are differently programmed by the presence or absence of *en* function, they respond in different ways to *dpp*, generating distinct anterior (A) and posterior patterns (P). The dotted lines indicate the position of the dorsal-ventral compartment boundaries in both discs. The boundary in the wing disc is well documented, whereas the boundary in the leg disc is less certain. We suggest that the different responses of dorsal and ventral anterior cells in the leg disc to *hh* protein (expression of *dpp* or *wg*, respectively) may be a consequence of their belonging to differently determined dorsal and ventral compartments. Further, we suggest that proximal-distal patterning in both discs is governed by the generation of at least two signals at the intersection between the anterior-posterior and dorsal-ventral compartment boundaries. In the leg disc, these different signals could be *dpp* and *wg*; in the wing disc, they might be *dpp* and one other signalling molecule generated along the dorsal-ventral compartment boundary.

anterior leg compartment can organize symmetrical anterior-ventral patterns (for example of sex comb teeth in the first leg) whereas ectopic *wg* expression in the posterior compartment organizes symmetric posterior patterns (for example, of transverse row bristles in the third leg)³⁴. In contrast, loss of endogenous *wg* function blocks ventral patterning in both compartments^{22,47,54}. Thus, in normal leg discs, the secretion of *wg* protein by anterior cells at the compartment boundary may be sufficient to organize ventral patterning in both compartments, without any further involvement of *hh* protein.

If *wg* and *dpp* play equivalent roles in mediating the response to *hh* protein, then we would predict that ectopic *dpp* activity in both the wings and the dorsal leg should similarly organize

symmetrical anterior and posterior patterns depending on the compartmental provenance of the responding cells. For example, ectopic *dpp*-expressing cells in the wing, which normally displays a 123/45 pattern of veins, might organize 123y321 patterns in the anterior wing compartment and 54y45 patterns in the posterior compartment. Thus, just as the juxtaposition between differently programmed cells along the compartment boundary leads to the asymmetric induction of *dpp* and *wg* expression, the different states of these cells may cause them to respond in different ways to *dpp* and *wg* protein emanating from the boundary, thereby generating distinct anterior and posterior patterns.

Range of action of *hh* protein

In the embryonic ectoderm, *hh*, like *en*, is expressed solely in posterior cells^{20,41-43} and appears to be required only in adjacent anterior cells to sustain expression of *wg*^{28,32}. Further, ectopic *hh* expression causes more anteriorly situated cells to express *wg*³¹. These data have been interpreted as evidence for local action of *hh*, possibly requiring contact between *hh*-expressing and *hh*-responding cells^{28,31,32}.

Our findings suggest that *hh* protein may act over a longer distance in the imaginal discs. For example, the stripe of *dpp* expression which runs along the anterior-posterior compartment in the wing disc is several cells wide^{27,44-46}. Because all anterior wing cells can respond to ectopic *hh* activity by expressing *dpp* (Fig. 4), we infer that the width of the *dpp* stripe reflects the upper limit of the range over which *hh* protein secreted by posterior cells can directly interact with and influence the behaviour of anterior cells. The same argument applies to *dpp* and *wg* expression in the leg discs. Thus, *hh* protein secreted by posterior cells may spread to and directly influence the behaviour of anterior cells at a distance of several cell diameters from the compartment boundary.

Responses to *hh* protein by different cells

Although all anterior cells appear responsive to spatially indiscriminate *hh* protein, we find that the nature of the response (high levels of *dpp* or *wg* expression) can depend on disc type and relative position within the disc. In the wing disc, all anterior cells can respond by expressing *dpp*. In contrast, *wg* expression in these cells, which correlates with the segregation of dorsal-ventral rather than anterior-posterior compartments⁵⁵, does not appear to be altered by either gain or loss of *hh* function (Fig. 5c, d). In the leg disc, dorsally situated cells respond by secreting *dpp* whereas ventrally situated cells secrete *wg*. The basis of this differential response is uncertain. However, it is notable that the *dpp*- and *wg*-responding cell populations in the leg disc seem to be complementary (Fig. 4d, f). Moreover, they may coincide with a possible dorsal-ventral compartmental segregation in the anterior compartment⁵⁶. Thus, all anterior cells may receive the *hh* signal through the same signal transduction mechanism, but distinct subpopulations may be programmed to mediate the organizing activity of *hh* by secreting different classes of signalling molecules.

Patterning along the proximal-distal axis

Our observation that leg discs that contain ectopic *hh*-expressing cells in both dorsal (*dpp*-responding) and ventral (*wg*-responding) regions can form supernumerary limbs (Fig. 3) supports the view⁵⁷ that the juxtaposition of *dpp*-secreting and *wg*-secreting cells may suffice to direct development along the proximal-distal axis. Similarly, we find that wing discs that contain ectopic *hh*-expressing cells in the vicinity of the prospective wing margin can form supernumerary appendages (Fig. 2b). In this case, the wing margin is known to coincide with a dorsal-ventral compartment boundary which is defined both by cell lineage and the differential expression of a selector gene, *apterous* (*ap*)^{1,3,58,59}. Hence in both the wing and leg discs, supernumerary limbs may arise when neighbouring anterior cells on both sides of a dorsal-ventral compartment boundary are exposed to ectopic *hh* pro-

tein. We therefore infer that proximal-distal patterning may normally depend on the differential response of dorsal and ventral cells to *hh* protein. In the leg disc this would involve the secretion of *dpp* and *wg*, respectively. In the wing disc, however, all anterior cells respond to *hh* by expressing *dpp*. Consequently, additional signals required for organizing the proximal-distal axis may be generated at the boundary between dorsal (*ap-on*) and ventral (*ap-off*) cells, possibly by a mechanism analogous to that involving *en*, *hh* and *dpp* along the anterior-posterior compartment boundary⁵⁹.

Implications for *hh* homologues in vertebrates

hh homologues have recently been reported in several vertebrate systems and implicated in the organization of limb and neural tube development^{38–40}. Of particular relevance to our findings, one homologue, *Sonic-hedgehog* (*Shh*), is normally expressed in the posterior zone of polarizing activity (ZPA) of the chick limb and is capable of organizing duplicated limb patterns with reversed polarity when ectopically expressed on the opposite (anterior) side of the limb³⁸. Thus, both *hh* and *Shh* are normally expressed in posterior domains of the limb primordia; moreover their ectopic expression in anterior regions can organize symmetrical pattern duplications composed of more anterior pattern elements. Because these genes appear to be functionally³⁹ as well as structurally homologous, we note three implications of our findings for vertebrate systems.

First, our results suggest that vertebrate *hh* proteins will not organize limb and neural tube patterning by acting as gradient morphogens (by specifying distinct outputs at different concentrations), but rather act as inducers which exert their organizing influence indirectly by triggering the expression of TGF- β or *Wnt* related proteins, or perhaps other classes of signalling mol-

ecules. As we propose for *Drosophila* limbs, such secondary signalling molecules could emanate symmetrically from a source of *hh*-responsive cells that about *hh*-expressing cells and organize different patterns of growth and differentiation in each population.

Second, our results provide a clear precedent that distinct subpopulations of cells within a tissue can mediate the organizing activity of *hh* protein by secreting different classes of signalling molecules.

Third, our results raise the question of whether the organizing activity of *hh* proteins in vertebrates is linked to a compartmental mechanism such as we propose for *Drosophila* limbs. We note, however, that there is at least one situation in *Drosophila*, the development of the eye, in which *hh*-dependent patterning^{30,33} is not related to compartments. In this case, *hh* expression defines an interface between anteriorly situated cells which respond to ectopic *hh* activity by expressing *dpp* and posteriorly situated cells which are refractory (refs 30, 33 and our unpublished observations). This interface is not related in any way to cell ancestry⁶⁰. Similarly, in vertebrates, the domain of *Shh* expression in both the limb and floor plate is likely to depend on local induction from neighbouring tissues, rather than on the inherited commitment of these cells to express *hh* protein^{61–64}. Thus, the organizing activity of vertebrate *hh* proteins may not be linked to compartments. However, other aspects of *hh* function, particularly the regulatory circuitry that governs the secretion of other signalling molecules and the response of neighbouring tissues, may be conserved.

Note added in proof: Tabata and Kornberg⁷² provide independent evidence that *hh* protein secreted by posterior cells can induce neighbouring anterior cells to express *dpp* as well as other genes along the compartment boundary. □

Received 23 December 1993; accepted 18 February 1994.

- Garcia-Bellido, A., Ripoll, P. & Morata, G. *Nature New Biol.* **245**, 251–253 (1973).
- Crick, F. H. & Lawrence, P. A. *Science* **189**, 340–347 (1975).
- Garcia-Bellido, A. *Genetic Control of Wing Disc Development in Drosophila* (Associated Scientific Publishers, New York, 1975).
- Morata, G. & Lawrence, P. A. *Nature* **255**, 614–617 (1975).
- Lawrence, P. A. & Morata, G. *Dev. Biol.* **50**, 321–337 (1976).
- DiNardo, S., Kuner, J. M., Theis, J. & O'Farrell, P. H. *Cell* **43**, 59–69 (1985).
- Kornberg, T., Siden, I., O'Farrell, P. H. & Simon, M. *Cell* **40**, 45–53 (1985).
- Hama, C., Ali, Z. & Kornberg, T. B. *Genes Dev.* **4**, 1079–1093 (1990).
- Moazed, D. & O'Farrell, P. H. *Development* **116**, 805–810 (1992).
- Vincent, J. P. & O'Farrell, P. H. *Cell* **68**, 923–931 (1992).
- Lawrence, P. A. & Struhl, G. *EMBO J.* **1**, 827–833 (1982).
- Meinhardt, H. *Dev. Biol.* **96**, 375–385 (1983).
- Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795–801 (1980).
- Spencer, F. A., Hoffmann, F. M. & Gelbart, W. M. *Cell* **28**, 451–461 (1982).
- Sharma, R. P. & Chopra, V. L. *Dev. Biol.* **48**, 461–465 (1976).
- van den Heuvel, M., Nusse, R., Johnston, P. & Lawrence, P. A. *Cell* **59**, 739–749 (1989).
- Panganiban, G. E., Rashka, K. E., Netzel, M. D. & Hoffmann, F. M. *Molec. cell. Biol.* **10**, 2669–2677 (1990).
- Panganiban, G. E., Reuter, R., Scott, M. P. & Hoffmann, F. M. *Development* **110**, 1041–1050 (1990).
- Gonzalez, F., Swales, L., Bejsovec, A., Skaer, H. & Martinez-Arias, A. *Mech. Dev.* **35**, 43–54 (1991).
- Lee, J. J., von Kessler, D. P., Parks, S. & Beachy, P. A. *Cell* **71**, 33–50 (1992).
- Taylor, A. M., Nakano, Y., Mohler, J. & Ingham, P. W. *Mech. Dev.* **43**, 89–96 (1993).
- Baker, N. E. *Dev. Biol.* **125**, 96–108 (1988).
- DiNardo, S., Sher, E., Heemskerk, J. J., Kassis, J. A. & O'Farrell, P. H. *Nature* **332**, 604–609 (1988).
- Martinez-Arias, A., Baker, N. E. & Ingham, P. W. *Development* **103**, 157–170 (1988).
- Mohler, J. *Genetics* **120**, 1061–1072 (1988).
- Immerglück, K., Lawrence, P. A. & Bienz, M. *Cell* **62**, 261–268 (1990).
- Posakony, L. G., Raftery, L. A. & Gelbart, W. M. *Mech. Dev.* **33**, 69–82 (1990).
- Ingham, P. W., Taylor, A. M. & Nakano, Y. *Nature* **353**, 184–187 (1991).
- Ferguson, E. L. & Anderson, K. V. *Cell* **71**, 451–461 (1992).
- Heberlein, U., Wolff, T. & Rubin, G. M. *Cell* **75**, 913–926 (1993).
- Ingham, P. W. *Nature* **366**, 560–562 (1993).
- Ingham, P. W. & Hidalgo, A. *Development* **117**, 283–291 (1993).
- Ma, C., Zhou, Y., Beachy, P. A. & Moses, K. *Cell* **75**, 927–938 (1993).
- Struhl, G. & Basler, K. *Cell* **72**, 527–540 (1993).
- Cabrera, C. V., Alonso, M. C., Johnston, P., Phillips, R. G. & Lawrence, P. A. *Cell* **50**, 659–663 (1987).
- Padgett, R. W., St Johnston, R. D. & Gelbart, W. M. *Nature* **325**, 81–84 (1987).
- Rijsewijk, F. et al. *Cell* **50**, 649–657 (1987).
- Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. *Cell* **76**, 1401–1416 (1993).
- Krauss, S., Concordet, J.-P. & Ingham, P. W. *Cell* **76**, 1431–1444 (1993).
- Echelard, Y. et al. *Cell* **76**, 1417–1430 (1993).
- Mohler, J. & Vani, K. *Development* **115**, 957–971 (1992).
- Tabata, T., Eaton, S. & Kornberg, T. B. *Genes Dev.* **6**, 2635–2645 (1992).
- Tashiro, S. et al. *Gene* **124**, 183–189 (1993).
- Masucci, J. D., Miltenberger, R. J. & Hoffmann, F. M. *Genes Dev.* **4**, 2011–2023 (1990).
- Blackman, R. K., Sanicola, M., Raftery, L. A., Gillevet, T. & Gelbart, W. M. *Development* **111**, 657–666 (1991).
- Raftery, L. A., Sanicola, M., Blackman, R. K. & Gelbart, W. M. *Development* **113**, 27–33 (1991).
- Couso, J. P., Bate, M. & Martinez-Arias, A. *Science* **259**, 484–489 (1993).
- Broach, J. R. & Hicks, J. B. *Cell* **21**, 501–508 (1980).
- Theurkauf, W. E., Baum, H., Bo, J. & Wensink, P. C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8477–8481 (1986).
- Kassis, J. A., Noll, E., VanSickle, E. P., Odenwald, W. F. & Perrimon, N. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1919–1923 (1992).
- Lawrence, P. A., Johnston, P., Macdonald, P. & Struhl, G. *Nature* **328**, 440–442 (1987).
- Bejsovec, A. & Martinez-Arias, A. *Development* **113**, 471–485 (1991).
- Heemskerk, J., DiNardo, S., Kostriken, R. & O'Farrell, P. H. *Nature* **352**, 404–410 (1991).
- Peifer, M., Rauskolb, C., Williams, M., Riggelman, B. & Wieschaus, E. *Development* **11**, 1029–1043 (1991).
- Williams, J. A., Paddock, S. W. & Carroll, S. B. *Development* **117**, 571–584 (1993).
- Steiner, E. *Wilhelm Roux Arch. EntwMech. Org.* **180**, 9–30 (1976).
- Campbell, G., Weaver, T. & Tomlinson, A. *Cell* **74**, 1113–1123 (1993).
- Blair, S. S. *Development* **119**, 339–351 (1993).
- Diaz-Benjumea, F. J. & Cohen, S. M. *Cell* **75**, 741–752 (1993).
- Ready, D. F., Hanson, T. E. & Benzer, S. *Dev. Biol.* **53**, 217–240 (1976).
- Anderson, R., Landry, M. & Muneoka, K. *Development* **117**, 1421–1433 (1993).
- Niswander, L., Tickle, C., Vogel, A., Booth, I. & Martin, G. R. *Cell* **75**, 579–587 (1993).
- Placzek, M., Jessell, T. M. & Dodd, J. *Development* **117**, 205–218 (1993).
- Vogel, A. & Tickle, C. *Development* **119**, 199–206 (1993).
- Struhl, G. *Dev. Biol.* **84**, 372–385 (1981).
- Golic, K. *Science* **252**, 958–961 (1991).
- Xu, T. & Rubin, G. M. *Development* **117**, 1223–1237 (1993).
- Morata, G. & Ripoll, P. *Dev. Biol.* **42**, 211–221 (1975).
- Patel, N. H. et al. *Cell* **58**, 955–968 (1989).
- Eaton, S. & Kornberg, T. B. *Genes Dev.* **4**, 1068–1077 (1990).
- Orenic, T. V., Slusarski, D. C., Kroll, K. L. & Holmgren, R. A. *Genes Dev.* **4**, 1053–1067 (1990).
- Tabata, T. & Kornberg, T. *Cell* **76**, 89–102 (1994).

ACKNOWLEDGEMENTS. We thank M. Keggenhoff, A. Nakanishi and R. Perez for technical assistance, B. Holmgren for the *Tuba1* DNA and *cP-lacZ* flies, U. Heberlein and K. Moses for sending *FRT hh* and *hh²²* stocks before publication, R. Blackman and W. Gelbart for the *dpp-lacZ* stock, M. Hoffman for the *dpp* antibody, M. van der Heuvel for the *wg* antibody, T. Kornberg for exchanging information before publication, and G. Campbell, I. Greenwald, E. Hafen, T. Jessell, J. Jiong, P. A. Lawrence, R. Mann and A. Tomlinson for discussion and critical comments on the manuscript. This work was supported by the Howard Hughes Medical Institute and the Swiss National Science Foundation.

Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide

Lawrence J. Stern^{*}, Jerry H. Brown[†], Theodore S. Jardetzky^{*},
Joan C. Gorga[‡], Robert G. Urban, Jack L. Strominger & Don C. Wiley^{*§}

Department of Biochemistry and Molecular Biology, and the ^{*} Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

[†] Rosenstiel Research Center, Brandeis University, 415 South Street, Waltham, Massachusetts 02254, USA

[‡] Rangos Research Center, Children's Hospital of Pittsburgh and University of Pittsburgh Medical Center, 3705 Fifth Avenue, Pittsburgh, Pennsylvania 15213, USA

An influenza virus peptide binds to HLA-DR1 in an extended conformation with a pronounced twist. Thirty-five per cent of the peptide surface is accessible to solvent and potentially available for interaction with the antigen receptor on T cells. Pockets in the peptide-binding site accommodate five of the thirteen side chains of the bound peptide, and explain the peptide specificity of HLA-DR1. Twelve hydrogen bonds between conserved HLA-DR1 residues and the main chain of the peptide provide a universal mode of peptide binding, distinct from the strategy used by class I histocompatibility proteins.

HISTOCOMPATIBILITY proteins are cell-surface glycoproteins that bind peptides within the cell and present them at the cell surface for interaction with T cells as part of the immune system's mechanism for identifying and responding to foreign antigens (reviewed in refs 1 and 2). Engagement of a histocompatibility protein with its bound peptide by the antigen receptor on a T cell causes stimulation of the T cell and activation of an immune response. Each individual expresses only a small number of different histocompatibility proteins. Thus each histocompatibility protein must be able to bind a large number of different peptides in order to ensure an immune response against many possible pathogens. The extensive polymorphism of histocompatibility genes (more than 50 alleles at some loci) may be the result of selection of alleles that can present peptides from particular pathogens.

Class I and class II histocompatibility proteins have different domain organizations but similar structures, with two membrane-proximal immunoglobulin-like domains and a membrane-distal peptide-binding site formed by an eight-stranded β -sheet and two α -helical regions^{3,4}. Polymorphic residues in both class I and II proteins are clustered in the peptide-binding region and are responsible for the different peptide specificities observed for different histocompatibility proteins. Class I histocompatibility proteins are specific for peptides of defined length, usually 8–10 residues^{5–7}, and have allele-specific binding motifs characterized by strong preferences for a few side chains at some positions in the peptide, and wide tolerance for many side chains at the other positions⁵. Class II histocompatibility proteins bind longer peptides with no apparent restriction on peptide length^{8,9}. Class II proteins also have allele-specific motifs^{10,11}, which have been more difficult to characterize because of the difficulty in aligning peptide sequences of different lengths.

Crystallographic studies of several human^{12–15} and mouse^{16–18} class I histocompatibility proteins have identified a common mechanism for peptide binding to class I proteins. Peptides of different sequences all bind similarly, with their N and C termini sequestered in the binding site by a network of hydrogen bonds to residues conserved in all class I histocompatibility proteins¹⁹. Peptides of different lengths (8–11 residues) are generally accom-

modated by buckling in the middle^{15, 18, 20}. Peptide sequence specificity is provided by a small number of allele-specific pockets, which bind side chains from the bound peptide^{20, 21}. The mechanism of peptide binding to class II histocompatibility proteins has not been as clear. The 3.3 Å crystal structure of the human class II histocompatibility protein HLA-DR1 (ref. 4) showed that bound peptide extended out the ends of the binding site, but interpretation of HLA-peptide interactions was complicated by the presence of a mixture of endogenous peptides in the peptide-binding site.

We have crystallized HLA-DR1 with a single peptide, the well characterized antigenic peptide HA306–318 (refs 22, 23) from influenza virus haemagglutinin (HA), in order to interpret in detail the interactions in the peptide-binding site. The refined 2.75 Å structure of the HLA-DR1/HA peptide complex shows that the peptide binds as a straight extended chain with a pronounced twist. This conformation allows most of the peptide residues to participate in the surface presented to the T-cell receptor. Hydrogen bonds between main-chain atoms along the peptide and HLA-DR1 residues from the α -helical regions and the β -sheet provide a component to the binding interaction that is independent of peptide sequence. Twelve of the hydrogen bonds involve residues conserved in most human and mouse class II alleles, and suggest a universal method for peptide binding by class II histocompatibility proteins. Five side chains of the HA peptide are accommodated by polymorphic pockets in the HLA-DR1 binding site. These pockets appear to determine the peptide specificity of different class II proteins.

Structure determination

Soluble, recombinant HLA-DR1 (*DRA*, *DRB1*0101*), produced in insect cells and free of endogenous peptides²⁴, was used to prepare *in vitro* a 1:1 complex with a synthetic antigenic peptide corresponding to influenza haemagglutinin residues 306–318 (PKYVKQNTLKLAT). Diffraction data (Table 1) comprising 94% of the reflections between 20–2.75 Å were collected from a single crystal at –170 °C using synchrotron radiation. Initial phases were determined by molecular replacement using a 4.2 Å model of the HLA-DR1 ($\alpha\beta$)₂ dimer⁴ which did not include peptide coordinates. The resolution was extended to 2.75 Å by cycles of iterative non-crystallographic real-space

§ To whom correspondence should be addressed.

TABLE 1 Data collection and refinement statistics

Data collection statistics					
Number of measurements		88,058	Real space fit§	0.85 ± 0.05	0.84 ± 0.09
Number of independent reflections		29,857		(main)	(side)
Resolution range	20–2.75 Å	3–2.75 Å	Average atomic B factor	32 ± 13 Å ²	35 ± 17 Å ²
Completeness	94.5%	89.8%		(Cα only)	(all atoms)
R _{merge} [*]	0.067	0.231	Deviations from ideal geometry:	0.019 Å	3.3°
				(bonds)	(angles)
Reflections with I > 2σ _I	90.5%	79.6%			
Refinement statistics (7–2.75 Å)			Non-crystallographic symmetry		
Number of reflections in refinement (F > 2σ _F)		23,847	Euler angles	Translation	r.m.s. dev.
Number of atoms included in refinement†		6,604			
R _{free} ‡ 0.306			α ₁ /β ₁	69.9°, 149.1°, 175.0°	−0.55 Å
			α ₂	69.7°, 148.9°, 178.4°	−0.53 Å
R _{cryst} ‡ 0.205			β ₂	68.4°, 149.2°, 179.0°	−0.02 Å

Crystals in the tetragonal space group $P4_32_12(a=94.6, c=245.7 \text{ Å})$, two molecules per asymmetric unit) were grown as described²⁴. Diffraction data were collected from a single crystal at -170°C , using the F1 beamline ($\lambda=0.910 \text{ Å}$) at the Cornell High Energy Synchrotron Source, and processed using DENZO (Z. Otwinowski, personal communication) and the CCP4 package (SERC Collaborative Computing Project, Daresbury Laboratory). A 4.2 Å model of the HLA-DR1 ($\alpha\beta$)₂ dimer⁴ was positioned in the asymmetric unit by molecular replacement using MERLOT⁴⁴. The top two rotation function solutions (4.9σ and 4.5σ ; the next highest peak was 3.9σ) were related by 180° and identified the non-crystallographic dimer axis. The top translation function solution (9.6σ with no other peaks above 6σ) identified the correct space group enantiomorph ($P4_32_12$). After rigid body refinement of the non-crystallographic symmetry operators, R_{cryst} for this model was 42.7% ($9-5 \text{ Å}$). The model was rebuilt piecewise into $2F_{\text{obs}} - F_{\text{calc}}$ maps, each calculated by omitting approximately one-eighth of the model and improved by iterative real-space averaging⁴⁵ across the non-crystallographic twofold symmetry axis and by solvent flattening. Clear electron density corresponding to the bound HA peptide was observed. The model was improved by 15 cycles of iterative real-space non-crystallographic averaging and solvent flattening⁴⁵, model building using 'O'⁴⁶, and refinement using XPLOR⁴⁷, as the resolution was gradually increased to 2.75 Å . Refinement protocols usually included rigid-body, positional and restrained B-factor minimization, and occasionally simulated annealing. Simulated annealing omit maps⁴⁷ were employed to reduce model bias in the electron density maps. Restraining terms relating the domains to their non-crystallographically related partners were retained, with decreasing weights, throughout the refinement. After 6 cycles, an overall anisotropic B-factor ($B_{33}=8 \text{ Å}^2$) was applied, and after 9 cycles, restrained individual B-factors were refined. Solvent molecules were included after a cycle of geometric regularization. Criteria for inclusion of water molecules included: $F_{\text{obs}} - F_{\text{calc}}$ density $> 3.5\sigma$, standard hydrogen-bonding geometry, stable positional refinement, atomic B-factor $< 60 \text{ Å}^2$, and real-space fit values > 0.4 . At each cycle of refinement, only refinement strategies that reduced the statistically independent R_{free} parameter were chosen⁴⁷. The traditional crystallographic R-factor did not appear to be as reliable an indicator of model quality as the free R-factor. For example, at cycle 14, models obtained by different refinement protocols had crystallographic R-factors ranging from 0.209 to 0.234 but identical free R-factors of 0.309. These models were indistinguishable by visual inspection, and produced maps of similar quality. Conversely, models with higher free R-factors produced maps of lower quality. The final 2.75 Å $2F_{\text{obs}} - F_{\text{calc}}$ map (Fig. 1c) shows clear electron density for 376 of 381 residues (HLA-DR1 $\alpha 3-182$, $\beta 3-190$, HA 306–318), with only partial density for the side chains of 5 surface residues (Arg $\alpha 50$, Glu $\alpha 55$, Glu $\beta 59$, Glu $\beta 187$, Arg $\beta 189$). Electron density corresponding to the N-linked carbohydrates at Asn $\alpha 78$, Asn $\alpha 118$ and Asn $\beta 19$ was observed, but only one ($\alpha 78$ and $\beta 19$) or two ($\alpha 118$) monosaccharide units could be reliably modelled at each position. C-terminal residues $\alpha 183-192$ and $\beta 191-198$, which correspond to the 'connecting peptide' region of HLA-DR1 (exon 4), show weak density and were not modelled. In the final model, no consecutive residues have main-chain real-space fit values of less than 0.64. Three residues have (ϕ, ψ) angles that fall outside the allowed regions of the Ramachandran plot: Glu $\alpha 172$, Asn $\beta 33$ and Lys $\beta 105$. Regions deemed ambiguous in the averaged 3.3 Å structure⁴ were resolved, with residues $\beta 64-70$ requiring the most extensive rebuilding. Estimated coordinate error (Luzzatti plot⁵¹) is 0.4 Å . Thermal β -factors and real-space fit values for all residues of the HA peptide are within 1.5σ of the values for HLA-DR1, except for Thr 318, which does not contact HLA-DR1 and is likely to be disordered. Coordinates will be deposited with the Protein Data Bank (Brookhaven National Laboratory, Upton, NY)

* For R_{merge} , sum taken over multiply recorded reflections; $R_{\text{merge}} = \sum_{hkl} \sum_{\text{obs}} |I_{\text{obs}}^{hkl} - \langle I_{\text{obs}}^{hkl} \rangle| / \sum_{hkl} \sum_{\text{obs}} I_{\text{obs}}^{hkl}$

† Includes 376 amino-acid and four sugar residues per monomer, and 153 water molecules.

‡ For R_{cryst} , sum taken over all reflections in the working set (90% of all reflections, used for refinement); for R_{free} , sum taken over all reflections in the free set (10% of all reflections, excluded from refinement). $R_{\text{cryst, free}} = \sum_{hkl: F > 2\sigma} \|F_{\text{obs}} - F_{\text{calc}}\| / \sum_{hkl: F > 2\sigma} \|F_{\text{obs}}\|$

§ Real-space fits of model to $2F_{\text{obs}} - F_{\text{calc}}$ map, using model phases, determined for main-chain and side-chain protein atoms⁴⁶.

|| Non-crystallographic symmetry operators were slightly different for three different domains in the protein. Operators and r.m.s. deviations from these operators are shown.

averaging and solvent flattening, model building, and automated refinement. Details of the structure determination are given in Table 1.

Structure of the bound peptide

The HA peptide is bound to HLA-DR1 as a straight extended strand with the peptide N and C termini projecting out of the ends of the binding site (Fig. 1a, b). The peptide has a pronounced twist, and successive side chains project from the peptide backbone about every 130° (Fig. 2a, b) in a conformation characteristic of the type II polyproline helix. This conformation directs many peptide side chains towards interactions with HLA-DR1, while at the same time allowing the T-cell receptor extensive interaction with the bound peptide (Fig. 2b, c).

HA-peptide side chains that project away from the binding site (Lys 307, Val 309, Lys 310, Asn 312, Lys 315, and Thr 318) are spaced approximately every three residues, in the pattern

●↑●↑●↑●↑●↑●↑ (where ↑ represents a side chain oriented away from the binding site). All of these residues except Thr 318 at the C terminus have been implicated by previous experiments in interactions with the T-cell receptor²⁵. Between these are side chains that project towards or across the peptide-binding site (Pro 306, Tyr 308, Gln 311, Thr 313, Leu 314, Leu 316, and Ala 317). All of these residues except Pro 306 and Ala 317 at the ends of the peptide project into pockets in the peptide-binding site, and have been implicated previously in interactions with HLA-DR1 (refs 22, 23, 26). The extreme N- and C-terminal residues of HA306–318 do not appear to be critical for either HLA-DR1 binding²⁷ or T-cell interaction²⁵.

Every residue in the HA peptide except Thr 318 contacts HLA-DR1. Thirty-nine HLA-DR1 residues contact the bound peptide (Fig. 2a, d). The interaction between the peptide and HLA-DR1 buries $1,400 \text{ Å}^2$ of surface area accessible to solvent in the free peptide, and 930 Å^2 of surface area accessible to

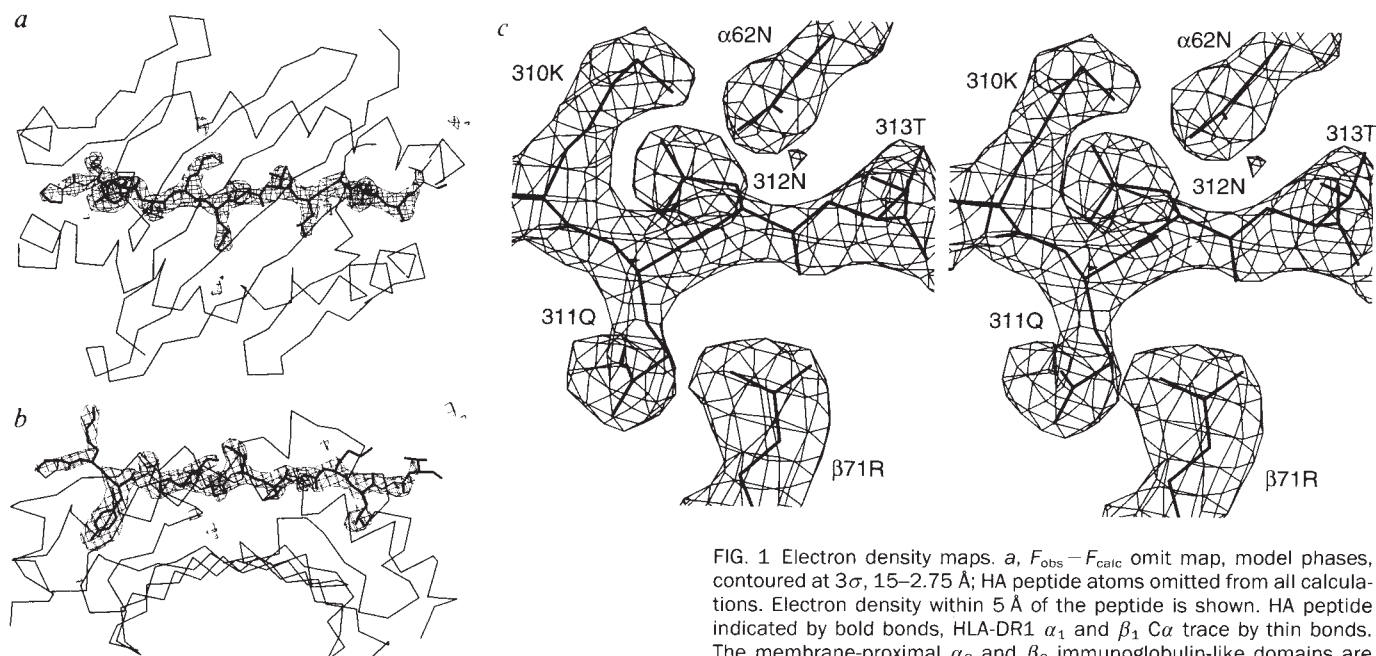


FIG. 1 Electron density maps. *a*, $F_{\text{obs}} - F_{\text{calc}}$ omit map, model phases, contoured at 3σ , 15–2.75 Å; HA peptide atoms omitted from all calculations. Electron density within 5 Å of the peptide is shown. HA peptide indicated by bold bonds, HLA-DR1 α_1 and β_1 Ca trace by thin bonds. The membrane-proximal α_2 and β_2 immunoglobulin-like domains are not shown. Top view shows α_1 helical region at the top, β_1 helical region at the bottom, HA peptide in the centre, with the N terminus on the left. *b*, Side view, 90° rotated from *a*, with β_1 helical region in front. *c*, Stereo view of region around HA peptide Asn 312 in $2F_{\text{obs}} - F_{\text{calc}}$ omit map, model phases, contoured at 1σ , 15–2.75 Å, with all atoms shown omitted from calculation; top view. Figure generated using 'O'46.

solvent in the unliganded HLA-DR1 (calculated using a probe radius of 1.4 Å; ref. 28). Approximately 35% of the peptide surface (750 Å²) remains solvent-accessible in the complex and potentially available for interactions with the T-cell receptor.

Hydrogen bonds with the peptide main chain

There are fifteen hydrogen bonds between HLA-DR1 and main-chain atoms of the bound HA 13-amino-acid peptide (Fig. 3*a*). Twelve of the fifteen hydrogen bonds involve HLA-DR1 atoms (Fig. 3*b*) that are conserved in most human and mouse class II proteins²⁹. HLA-DR1 main-chain atoms from the short extended strand at the N terminus of the α_1 helical region form three hydrogen bonds to the first two N-terminal residues of the bound peptide, in the same pattern of hydrogen bonding observed in parallel β -strands (Fig. 3, blue). Three conserved asparagine residues ($\alpha 62$, $\alpha 69$ and $\beta 82$) make bidentate hydrogen bonds to the peptide (Fig. 3, red). The interaction of the asparagine carbonyl oxygen and amide hydrogen atoms with similar atoms from the main chain of the peptide form 9-atom or 11-atom rings³⁰. Two conserved aromatic residues, Trp $\beta 61$ and His $\beta 81$ (Fig. 3, green), form hydrogen bonds with peptide carbonyl atoms. Trp $\beta 61$ is homologous to a conserved Trp ($\alpha 147$) in class I histocompatibility proteins that also binds a peptide carbonyl atom. Arg $\alpha 76$ (Fig. 3, green) and Asp $\beta 57$ hydrogen-bond with the main chain of the peptide, and form a salt bridge underneath the bound peptide which links the HLA-DR1 α_1 - and β_1 -chain helical regions. Arg $\alpha 76$ is conserved in class II α chains, and position $\beta 57$, although commonly aspartate, is polymorphic. Finally, two other polymorphic HLA-DR1 residues participate in hydrogen bonds with the main chain of the bound peptide, Gln $\alpha 9$ from the β -sheet and Arg $\beta 71$ from the β_1 -chain helical region.

The main chain of every peptide residue except Lys 310 and Thr 318 makes at least one hydrogen bond to HLA-DR1. The fifteen hydrogen bonds are spaced evenly along the peptide-binding site (Fig. 3*a*) and probably contribute substantially to the tight binding of HA306–318 to HLA-DR1. As they do not involve the side chains of the peptide, these hydrogen bonds are likely to play a universal role in the binding of other peptides to HLA-DR1. Moreover, twelve of the hydrogen bonds to the main chain of the peptide (Fig. 3*b*) involve HLA residues that are conserved in most other class II proteins. Other class II proteins are likely to use these same hydrogen bonds to bind peptides.

Side-chain pockets and peptide specificity

Within the peptide-binding site of HLA-DR1 (Fig. 4) are several smaller cavities, or 'pockets', five of which accommodate side chains of the bound HA peptide (Fig. 4). Many of the residues that line these pockets are highly polymorphic²⁹, and this polymorphism is probably responsible for the different peptide specificities of different class II proteins.

The side chain of the HA peptide Tyr 308 is buried in the largest (~200 Å² of contact surface area²⁸) and most hydrophobic pocket in the binding site (labelled '1' in Fig. 4*a*). This pocket shows a strong preference for large hydrophobic side chains (Trp, Tyr, Phe, Leu and Ile)^{26,31} and appears to be the major determinant of HA peptide binding to HLA-DR1 (ref. 22). Among different HLA-DR proteins, much of this pocket is conserved (all α -chain residues, and Phe $\beta 89$ and Thr $\beta 90$), and many other HLA-DR proteins also prefer a large hydrophobic peptide side chain at this position^{10,26}. Specificity in this pocket appears to be modulated by a Gly/Val dimorphism at position $\beta 86$ (refs 32,33) which changes the preference towards smaller side chains^{34,35}. The substitution of Gly $\beta 86$ by valine would place the valine side chain within the region occupied by the HA peptide tyrosine hydroxyl group, and would restrict the size of the peptide side chain that could be accommodated at this position.

Four other HA peptide side chains (Gln 311, Thr 313, Leu 314 and Leu 316) are more than 90% sequestered from solvent by pockets in the HLA-DR1 binding site (Fig. 4*a*). These pockets are smaller, more shallow, and less hydrophobic than the pocket that accommodates Tyr 308, and are more permissive in the side chains that they can accommodate^{26,31}. Together with the requirement for a large hydrophobic residue in position '1', the preferences in these four 'minor' pockets seem to determine most of the peptide specificity, or binding motif, of HLA-DR1.

The side chain of Gln 311 lies in a large, shallow pocket ('4' in Fig. 4*a*) oriented across the binding side, with the terminal amide partly exposed (Fig. 4*b*). This pocket can bind a variety of large aliphatic side chains³¹, which would retain the hydrophobic

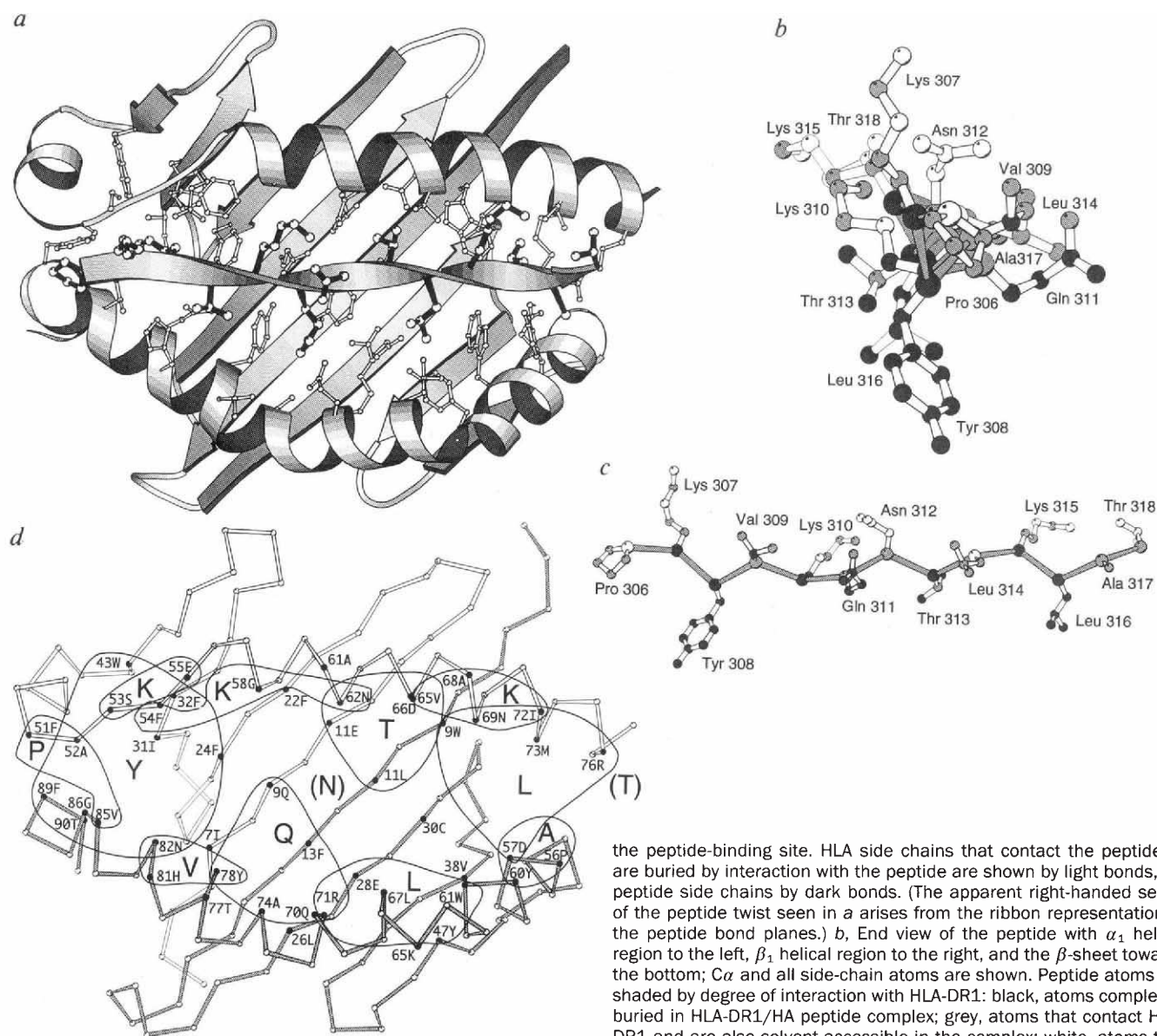


FIG. 2 HLA-DR1/HA peptide contacts. The side chains of the peptide project with a pronounced left-handed twist ($-130 \pm 15^\circ$ per residue, mean $\pm$ standard deviation). Peptide ϕ , ψ angles are $-82 \pm 15^\circ$, $139 \pm 16^\circ$. This conformation is similar to the polyproline type II helix (-120° per residue, $\phi = -78^\circ$, $\psi = 149^\circ$), and distinct from canonical parallel (ϕ about -139° , ψ about $+135^\circ$) or non-parallel (ϕ about -119° , ψ about $+113^\circ$) β -strands ($+160$ to $+180^\circ$ per residue), which usually have a slight right-handed twist. *a*, Ribbon diagram, top view of

the peptide-binding site. HLA side chains that contact the peptide or are buried by interaction with the peptide are shown by light bonds, HA peptide side chains by dark bonds. (The apparent right-handed sense of the peptide twist seen in *a* arises from the ribbon representation of the peptide bond planes.) *b*, End view of the peptide with α_1 helical region to the left, β_1 helical region to the right, and the β -sheet towards the bottom; $C\alpha$ and all side-chain atoms are shown. Peptide atoms are shaded by degree of interaction with HLA-DR1: black, atoms completely buried in HLA-DR1/HA peptide complex; grey, atoms that contact HLA-DR1 and are also solvent-accessible in the complex; white, atoms that do not contact HLA-DR1. *c*, As *b*, except that it is a side view. *d*, Map of HLA-DR1/HA peptide contacts; HLA-DR1 is indicated by the $C\alpha$ trace; α -chain light bonds; β -chain, dark bonds. Lines enclose all HLA-DR1 residues in contact (<4 Å) with each HA peptide side chain, labelled in the one-letter amino-acid code. Other HLA-DR1 residues that do not directly contact the peptide but are inaccessible to solvent owing to the interaction with the peptide, and which may directly contact the peptide in other class II/peptide complexes, are also indicated. HA side chains Asn 312 and Thr 318 do not contact HLA-DR1. Figure generated with Molscript⁴⁸.

interactions along the side and floor of the pocket. Positively charged residues are disfavoured at this position²⁶, presumably because of unfavourable electrostatic interactions with Arg 87. The side chain of Thr 313 is completely buried in a shallow pocket ('6' in Fig. 4a) between the α_1 -helical region and the β -sheet. Smaller residues are generally preferred at this position in other peptides¹⁰, and for the HA peptide the interaction in this pocket may be energetically unfavourable, as substitution of Thr 313 with alanine increases the binding affinity for HA306-318 (refs 22,31). The side chain of Leu 314 is bound in a shallow pocket ('7' in Fig. 4a) along the β -chain helix, but is only partially buried by the interaction (Fig. 4b, c). This pocket does not appear to contribute greatly to the peptide specificity of HLA-DR1, but may be important in other alleles. The side chain of Leu 316 is oriented down into a small hydrophobic pocket ('9' in

Fig. 4a). Hydrophobic residues are preferred at this position^{26,31}. Several other side chains of the HA peptide make appreciable contacts with HLA-DR1 (Fig. 1d), but are not bound in 'pockets'. These residues may nonetheless contribute to the binding through interactions with surface residues of HLA-DR1.

A great number of different peptide sequences can bind to HLA-DR1 with high affinity, because most of the pockets accommodate a wide variety of side chains, and because the binding reaction also includes sequence-independent contributions from interactions with the main chain of the peptide. Unfavourable interactions in one or more pockets can be compensated by favourable interactions elsewhere along the peptide. The overall specificity of peptide binding may be dominated by the inability to bind peptide side chains with poor electrostatic or van der Waals interactions, which cannot be accommodated

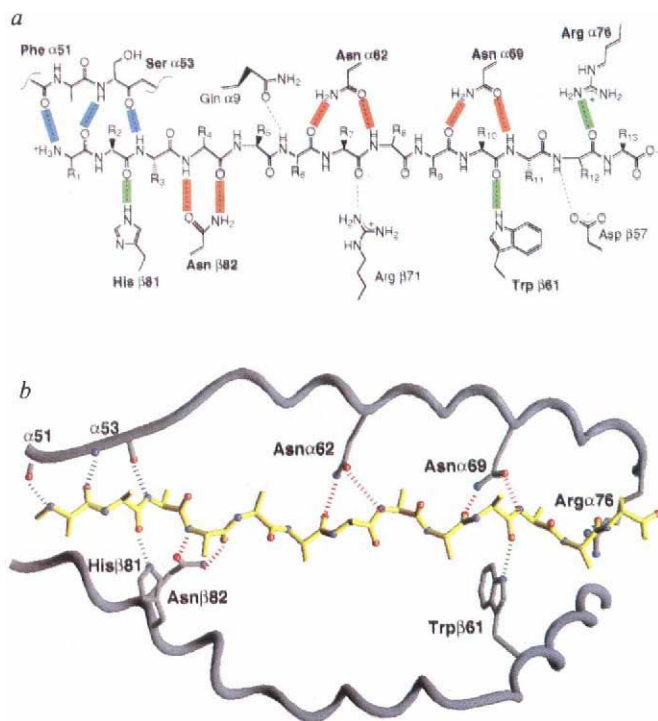


FIG. 3 Hydrogen bonds between HLA-DR1 and the HA peptide main chain. **a**, Top view: hydrogen bonds are indicated by dashed lines; conserved class II residues are shown in bold with hydrogen bonds coloured. Peptide side chains indicated by R₁-R₁₃; hatched and bold bonds indicate, respectively, peptide side chains that are oriented towards and away from the peptide-binding site, and do not indicate the stereochemistry of the peptide Cα atoms. The R₉ amide participates in an indirect hydrogen bond (not shown) to HLA-DR1 through a bound water molecule and the polymorphic residues Trp β9 and Asp β28. **b**, Top view, showing spatial arrangement of hydrogen bonds from conserved class II residues to the peptide main chain; hydrogen bonds coloured as in **a**. Figure generated with ChemDraw (Cambridge Scientific Computing) and Ribbons⁴⁹.

except by disrupting favourable main-chain or side-chain interactions elsewhere^{22,36}.

If the conformations of other peptides bound to class II molecules were similar to that of the HA peptide bound to HLA-DR1, then the pattern of side-chain contacts observed for the HA peptide (Fig. 2d) could be used to predict the binding specificity of other alleles. The correspondence between the interactions seen with the HA peptide and the HLA-DR1 binding motif suggests that many other peptides can bind to HLA-DR1 in the same manner as the HA peptide. Some motifs identified for other class II alleles can be interpreted in the framework of the HLA-DR1/HA peptide complex. For example, HLA-DRw17 (*DRB1*0301*), which has arginine instead of alanine at position β74, prefers aspartate or glutamate at the position that corresponds to HA Gln 311 (Fig. 4a)^{37,38}. The addition of a positive charge near the end of pocket 4 (Figs 2d, 4a) could explain the specificity for a negatively charged peptide side chain at this position. HLA-DR5 (*DRB1*1101*), which has aspartate instead of Trp β9, prefers positively charged side chains at the position that corresponds to Thr 313 in the HA peptide¹⁰. Although Trp β9 principally interacts with Leu 316 in the HLA-DR1/HA peptide complex (Figs 2d, 4a), it forms the boundary between pockets 6 and 9 and could interact with the side chain in pocket 6 with minimal rearrangement of the binding site. Although these cases suggest that many peptides bound to class II proteins share the conformation adopted by the HA peptide,

preliminary modelling experiments suggest that other peptide conformations may be possible (L.J.S., unpublished results). Moreover, in the HLA-DR1 binding site several polymorphic residues (Asp α66, Trp β9, Leu β11, Phe β13, Glu β28, Cys β30 and Val β38) line a shallow pocket (Fig. 2d, between 'T', 'L' and 'L'; Fig. 4a between '6' and '7'), which does not bind a side chain from the HA peptide, but instead is filled partially by two solvent molecules. The main chain of the HA peptide lies over this pocket, but a peptide bound in a different conformation could direct a side chain down into it. The possibility of alternative peptide conformations suggests caution in interpreting motifs found for other class II histocompatibility proteins strictly in terms of the HLA-DR1/HA peptide structure presented here.

Interactions with the T-cell receptor

The surface available to the T-cell receptor comprises atoms from both HLA-DR1 and the HA peptide (Fig. 4b, c). Some residues in the HLA-DR1/peptide complex that may interact with the T-cell receptor have been identified by mutagenesis studies. Substitutions of HA peptide residues that point away from the binding site (Lys 307, Val 309, Lys 310, Asn 312 and Lys 315; Figs 2c, 4c)^{25,23}, or of HLA-DR residues located near the pronounced kink in the β-chain helical region that forms the highest point of the structure (Leu β67, Gln β70 and Arg β71)^{39,40} (Figs 2a, 4c), alter recognition of the complex by various T-cell clones. These are all exposed residues, and their substitution would directly alter the surface of the HLA-DR1/peptide complex that interacts with the T-cell receptor.

Substitution of buried residues can also alter T-cell recognition. Substitution of residues that correspond to HLA-DR1 β30 (refs 36, 39) and β86 (refs 34, 41) or peptide residues that correspond to Tyr 308 (refs 35, 39), have all been shown to alter T-cell recognition with minimal effect on peptide binding. The location of these residues in buried regions suggests that their substitution must have an indirect effect, by altering the surface that directly contacts the T-cell receptor. In principle, any alteration in the binding site could be communicated through the peptide to the T-cell receptor.

Peptide binding compared for class I and II

Although the architecture of the peptide-binding site is very similar in class II and class I major histocompatibility complex (MHC) proteins (Fig. 5), they appear to have evolved to bind peptides very differently. Class II proteins bind peptides of apparently arbitrary length^{8,9} using conserved residues distributed throughout the binding site which make hydrogen bonds with the main chain of the bound peptide all along its length (Fig. 3b). Class I proteins generally bind shorter peptides⁵⁻⁷, using conserved residues at the ends of the binding site that make hydrogen bonds near the N and C termini of the bound peptide¹⁸⁻²⁰. In both class II and class I histocompatibility proteins, polymorphic pockets in the binding site provide the basis for the preference of different alleles for different peptide sequence motifs, but at least two more pockets accommodate peptide side chains in class II (about 5; Figs 2d and 4a) than in class I molecules (about 3; refs 18-20). The conformation of peptides bound to class II MHC molecules generally appears to be straight⁴ (Fig. 2), whereas peptides bound to class I MHC molecules are usually arched with a kink near P4 (Fig. 5b).

These differences in modes of peptide binding between class II and class I histocompatibility proteins result from a series of small structural alterations within a relatively fixed framework of the MHC binding site. At the N-terminal side of the class II binding site (Fig. 5, left), a short extended chain (α51-55; Fig. 2a, d) allows hydrogen bonding between main-chain HLA residues and the peptide (Fig. 3), and orients HLA side chains so that they do not sterically interfere with N-terminal extensions of the peptide. The large aromatic side chains that form the N-terminal end of the class I binding site (Tyr 159, Trp 167, Tyr 171) correspond to smaller class II residues (Ala β74,

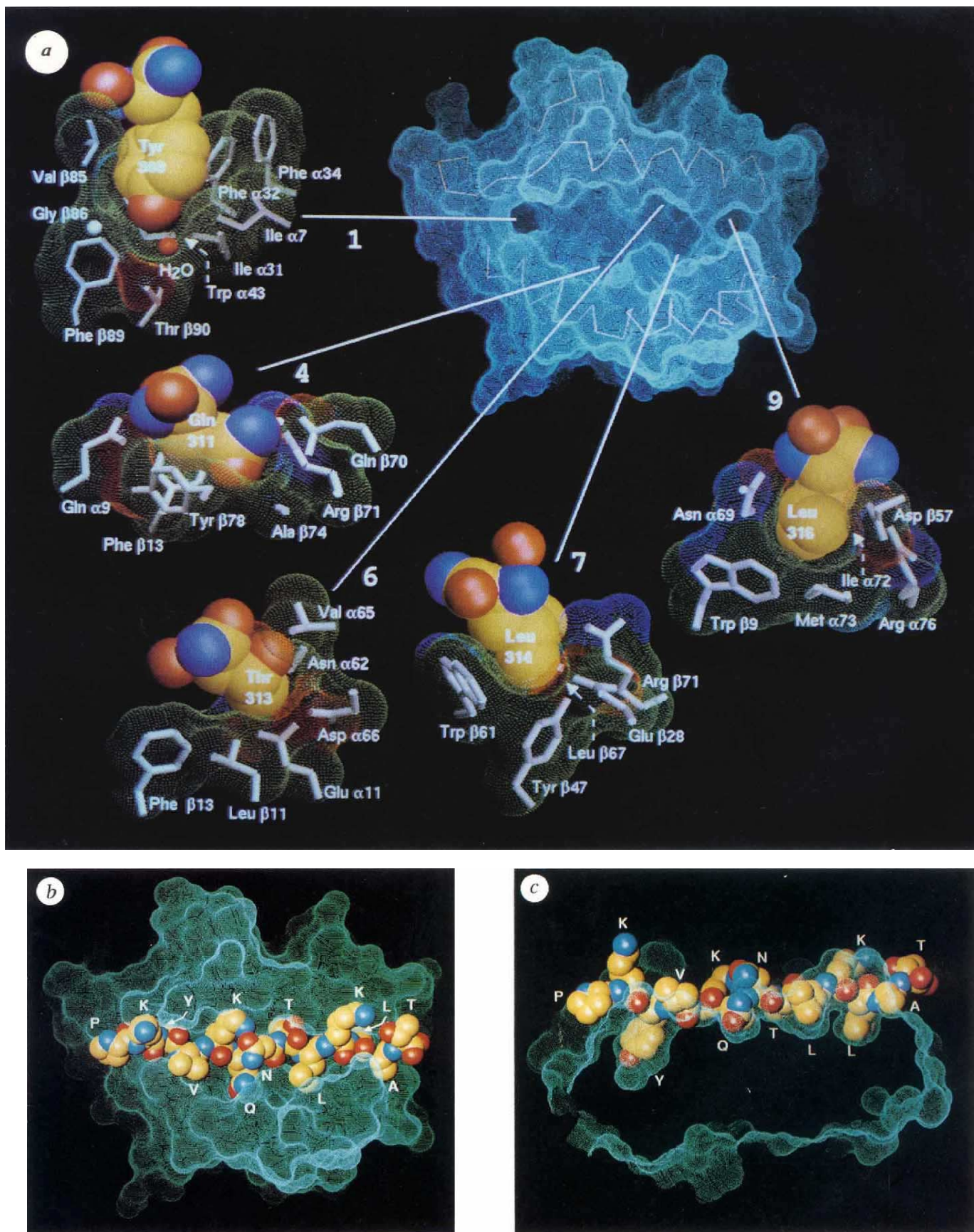
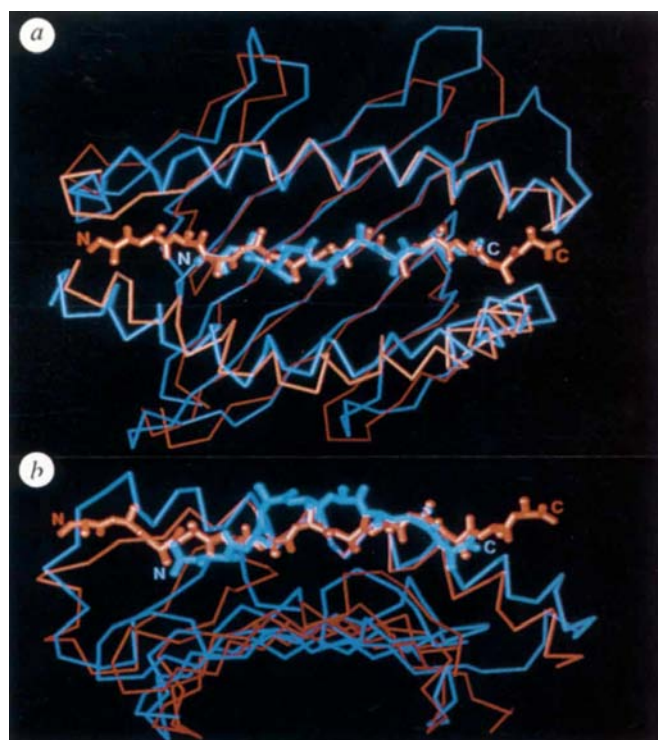


FIG. 4 Pockets in the HLA-DR1 peptide binding site. *a*, Upper right, top view of the molecular surface (1.5 Å probe radius)⁵⁰ of the HLA-DR1 peptide-binding site, shown in blue, with the C α trace of HLA-DR1 shown as thin yellow lines. Pockets that accommodate HA peptide side chains are shown in detail in surrounding views, and are numbered according to their distance along the peptide from the most prominent pocket (1) that accommodates the side chain of Tyr 308: counterclockwise from upper left, pockets that accommodate Tyr 308 (1), Gln 311 (4), Thr 313 (6), Leu 314 (7), Leu 316 (9). HA peptide side chains and the nearby peptide main chain are represented as CPK models; carbon atoms are yellow, nitrogen atoms blue, and

oxygen atoms red. HLA-DR1 side chains in contact with the peptide are indicated as white stick models. The molecular surface for HLA-DR1 atoms in each pocket is coloured by atom type: carbon and sulphur are green, oxygen is red, and nitrogen blue. Pockets are viewed in the plane of the peptide binding site, towards the N terminus of the peptide (1 and 6), towards the C terminus of the peptide (4), towards the β ₁ helical region (7) or towards the α ₁ helical region (9). *b*, Top view of the molecular surface of the HLA-DR1 peptide-binding site is shown in blue, with the HA peptide as a CPK model and peptide side chains labelled in one-letter code. *c*, As *b*, except from a side view. Figure generated with MS⁵⁰, 'O'⁴⁶, and Ribbons⁴⁹.



Asn β 82; Gly β 86) that leave the end of the class II site open. At the C-terminal end of the binding site, a slight difference in the orientation of the α_1 -domain α -helices (Fig. 5b, right) allows peptides bound to class II proteins to pass over the Arg α 76–Asp β 57 salt link, whereas peptides bound to class I proteins are often blocked at their C terminus by interaction with the homologous residues Tyr 84–Thr 143. (Some long peptides bound to class I proteins may extend out of the site⁴² (V. Englehard, personal communication), presumably by passing over Tyr 84–Thr 143.)

FIG. 5 Comparison of the peptide-binding sites of class I and class II HLA proteins. HLA binding sites shown as α traces, and the peptides as stick models with the side chains removed for clarity. Blue, class I HLA-B27 (ref. 13); red, class II HLA-DR1. a, Top view; b, side view, with α_2 (class I) and β_1 (class II) helical regions not shown for clarity. The r.m.s. deviation for peptide main-chain atoms between the 9-amino-acid peptide model in HLA-B27, and the corresponding portion of the HA-peptide (309–316) in HLA-DR1, is 3.1 Å. Figure generated with 'O'46.

Why have the two MHC structures become so differently adapted? Class II histocompatibility proteins bind peptides in endosomal or lysosomal compartments in which proteolytic processing occurs², and their capacity to bind long peptides that extend out of both ends of the site allows for the possibility that partial unfolding and proteolytic processing of antigen could occur while bound to the MHC molecule. Class I histocompatibility proteins bind peptides in the endoplasmic reticulum and may have evolved to bind short peptides through co-evolution with the proteolytic machinery that generates peptides in the cytoplasm, or with transporters that bring peptides into the endoplasmic reticulum².

Despite the difference in peptide binding between class I and class II MHC molecules, their peptide-binding affinity is similar, with the peptides that bind most tightly having nanomolar dissociation constants. The longer peptide-binding site of class II molecules^{4,43} comprises more HLA-peptide contact surface, offering the possibility of requiring less specificity at each contact. The peptide surface area available for direct recognition by a T-cell receptor (400–500 Å²) is about the same for several human class I-peptide complexes^{4,15} as for the central 11 residues of the HA peptide bound to class II HLA-DR1. The compound surface made up of peptides of different sequences, the parts of the MHC molecule that are left exposed by peptides of different sequences, and small distortions in the MHC molecule caused by peptide binding^{15,17}, offers direct stereochemical read-out of immunological foreignness. □

Received 15 November 1993; accepted 4 February 1994.

- Jorgensen, J. L., Reay, P. A., Ehrlich, E. W. & Davis, M. M. *A. Rev. Immun.* **10**, 835–873 (1992).
- Germain, R. N. & Margulies, D. H. *A. Rev. Immun.* **11**, 403–450 (1993).
- Bjorkman, P. J. et al. *Nature* **329**, 506–512 (1987).
- Brown, J. H. et al. *Nature* **364**, 33–39 (1993).
- Falk, K., Rötzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290–296 (1991).
- Hunt, D. F. et al. *Science* **255**, 1261–1263 (1992).
- Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R. & Wiley, D. C. *Nature* **253**, 326–329 (1991).
- Rudensky, A. Y., Preston-Hurlburt, P., Al-Ramadi, B. K., Rothbard, J. & Janeway, C. A. J. *Nature* **359**, 429–431 (1992).
- Chicz, R. M. et al. *Nature* **355**, 764–768 (1992).
- Hammer, J. et al. *Cell* **74**, 197–203 (1993).
- Sette, A. et al. *J. Immun.* **151**, 3163–3170 (1993).
- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Nature* **353**, 321–325 (1991).
- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Cell* **70**, 1035–1048 (1992).
- Silver, M. L., Guo, H.-C., Strominger, J. L. & Wiley, D. C. *Nature* **360**, 367–369 (1992).
- Madden, D. R., Garboczi, D. N. & Wiley, D. C. *Cell* **75**, 693–708 (1993).
- Zhang, W., Young, A. C. M., Imai, M., Nathenson, S. G. & Sacchettini, J. C. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8403–8407 (1992).
- Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. *Science* **257**, 919–927 (1992).
- Young, A. C. M., Zhang, W., Sacchettini, J. C. & Nathenson, S. G. *Cell* **76**, 39–50 (1994).
- Madden, D. R. & Wiley, D. C. *Curr. Opin. Struct. Biol.* **2**, 300–304 (1992).
- Matsumura, M., Fremont, D. H., Peterson, P. A. & Wilson, I. A. *Science* **257**, 927–934 (1992).
- Guo, H.-C. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8053–8057 (1993).
- Jardetzky, T. S. et al. *EMBO J.* **9**, 1797–1803 (1990).
- Rothbard, J. B. & Gefter, M. L. *A. Rev. Immun.* **9**, 527–565 (1991).
- Stern, L. J. & Wiley, D. C. *Cell* **68**, 465–477 (1992).
- Alexander, J. et al. *J. Immun.* **150**, 1–7 (1993).
- O'Sullivan, D. et al. *J. Immun.* **147**, 2663–2669 (1991).
- O'Sullivan, D., Sidney, J., del Guercio, M.-F., Colón, S. M. & Sette, A. *J. Immun.* **146**, 1240–1246 (1991).
- Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).
- Brown, J. H. et al. *Nature* **332**, 845–850 (1988).

- Le Questel, J.-Y., Morris, D. G., MacCallum, P. H., Peot, R. & Milner-White, E. J. *J. molec. Biol.* **231**, 888–896 (1993).
- Hammer, J., Takacs, B. & Sinigaglia, F. *J. exp. Med.* **176**, 1007–1013 (1992).
- Demotz, S., Barbey, C., Corradin, G., Amoroso, A. & Lanzavecchia, A. *Eur. J. Immun.* **23**, 425–432 (1993).
- Verreck, F. A. W., Termijtelen, A. & Konig, F. *Eur. J. Immun.* **23**, 1346–1350 (1992).
- Newton-Nash, D. K. & Eckels, D. D. *J. Immun.* **150**, 1813–1821 (1993).
- Busch, R., Hill, C. M., Hayball, J. D., Lamb, J. R. & Rothbard, J. B. *J. Immun.* **147**, 1292–1298 (1991).
- Boehncke, W.-H. et al. *J. Immun.* **150**, 331–341 (1993).
- Geluk, A. et al. *J. Immun.* **149**, 2864–2871 (1992).
- Malcherek, G. et al. *Int. Immun.* **5**, 1229–1237 (1993).
- Kreiger, J. I. et al. *J. Immun.* **146**, 2331–2340 (1991).
- Coppin, H. L. et al. *Eur. J. Immun.* **23**, 343–349 (1993).
- Ong, B. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7343–7347 (1991).
- Olsen, A. C. et al. *Eur. J. Immun.* (in the press).
- Driscoll, P. C. et al. *J. molec. Biol.* **232**, 342–350 (1993).
- Fitzgerald, P. M. D. *J. appl. Crystallogr.* **21**, 273–278 (1988).
- Bricogne, G. *Acta crystallogr.* **A32**, 832–847 (1976).
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110–119 (1991).
- Brünger, A. T. *X-PLOR (Version 3.1): A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).
- Kraulis, P. J. *appl. Crystallogr.* **24**, 946–950 (1991).
- Carson, M. J. *J. molec. Graphics* **5**, 103–106 (1987).
- Connolly, M. L. *J. appl. Crystallogr.* **16**, 548–558 (1983).
- Luzzati, V. *Acta crystallogr.* **5**, 802–811 (1952).

ACKNOWLEDGEMENTS. We thank D. Garboczi, S. Garman, P. Rosenthal, S. Watowich and the Cornell High Energy Synchrotron MacCHESS staff for assistance with data collection. S. Gamblin for discussions and M. Frayser for technical assistance in the preparation of HLA-DR1–HA peptide complexes; we also acknowledge allocation of computer resources from the Pittsburgh Supercomputer Facility. L.J.S. was supported by a Daymon Runyon–Walter Winchell fellowship; J.H.B. by an NIH predoctoral fellowship; T.S.J. by the Cancer Research Institute; J.G.C. by a Career Development Award from the Juvenile Diabetes Foundation International; R.G.U. by the Irvington Institute; J.L.S. by grants from the NIH; and L.J.S., J.H.B., and D.C.W. by the Howard Hughes Medical Institute.

Production of ^{26}Al and other extinct radionuclides by low-energy heavy cosmic rays in molecular clouds

Donald D. Clayton

Department of Physics and Astronomy, Clemson University, Clemson, South Carolina 29634-1911, USA

THE high abundance of radioactive ^{26}Al in meteoritic materials¹ presents a puzzle in attempts to describe the formation of the Solar System. It has been suggested¹¹ that collapse of the solar nebula may have been accompanied by an injection of ^{26}Al from nucleosynthesis in a neighbouring star. The relatively high level of ^{26}Al in the interstellar medium^{2,3} is also unexplained. Here I suggest that this isotope may be formed in nuclear reactions between hydrogen and heavy cosmic rays. This suggestion is prompted by the recent discovery⁴ of γ -ray line emission from ^{12}C and ^{16}O in the Orion cloud complex, thought to be caused by the interaction of these cosmic-ray ions with interstellar hydrogen. I show that these observations also imply enhanced production of ^{26}Al in the Orion molecular clouds, and that such nuclear reactions can account for the meteoritic abundances. Reactions involving heavy cosmic rays might also account for the presence of other extinct radioactive nuclides in meteorites.

The γ -ray line emission observed⁴ in the direction of the Orion molecular clouds has been attributed to the 4.43 MeV nuclear de-excitation of $^{12}\text{C}^*$ and the 6.13 MeV de-excitation of $^{16}\text{O}^*$. The excitations are proposed⁴ to happen when heavy cosmic rays interact with ambient H and He in the molecular clouds. The Solar-System cosmic-ray flux is too low (by a factor of 30) (ref. 5) to produce the observed inelastic excitation rate of $^{12}\text{C}^*$ and $^{16}\text{O}^*$ excited states, although the interstellar flux at energies as low as 10 MeV per nucleon is not known from observations in the Solar System⁶. The Compton telescope team⁴ therefore suggests heavy cosmic rays are 30 times more abundant in Orion. About two-thirds of the excess γ -ray emission seems to come from $^{12}\text{C}^*$ (4.43 MeV), and the remainder from $^{16}\text{O}^*$ (6.13 MeV). Their combined flux of $\sim 1 \times 10^{-4} \text{ cm}^{-2} \text{ s}^{-1}$ has been interpreted⁴ as arising from a region of $\sim 10^5$ solar masses ($M_\odot$), about 500 pc away. Cosmic-ray protons having standard galactic flux strike interstellar ^{12}C and ^{16}O at an inadequate rate⁵ for this γ -excitation rate; but the collision rate is much larger for carbon-ion cosmic rays striking interstellar H, because of the overabundance of carbon with respect to hydrogen in cosmic radiation⁶. For the subsequent discussion, I accept the conclusion from ref. 1 of a 30-fold enhancement in the cosmic-ray flux in Orion (specifically, in the heavy-nuclei cosmic rays with energies near 10 MeV per nucleon), and calculate other interactions proportionately. This enhancement implies that prolific acceleration of heavy cosmic rays to energies of 10–30 MeV per nucleon takes place in Orion, these cosmic rays interacting with and stopping in the molecular clouds. The reason for this preferential acceleration of heavy nuclei is not obvious, but might be related to the molecular state of H_2 in clouds.

^{26}Al can be produced from the reactions of hydrogen with Mg and Si in the cosmic rays. I assume that the abundances of Mg–Si relative to C–O are the same in Orion as those measured⁶ in the cosmic radiation near Earth, although it is not obvious that this should be the case. That Mg–Si/C–O ratio is also known to be enhanced relative to solar abundances by a factor of three (ref. 6) and the important ^{26}Mg abundance is isotopically enhanced by yet another factor of two (ref. 6) further enhancing the important collision rate for $^{26}\text{Mg} + \text{H}$.

The most important ^{26}Al -producing reactions^{7,8} are shown in Table 1. The first is especially important if energies extend down

to 3 MeV per nucleon, as has previously been considered in the solar disk itself⁸. In Table 1, I also compare the calculated ^{26}Al production rate with the observed $^{12}\text{C}^*$ excitation rate. The latter rate is $2 \times 10^{39} \text{ s}^{-1}$ in Orion clouds if they are 500 pc away, and the cross-section for production of $^{12}\text{C}^*$ (4.43 MeV) is 200 mb (ref. 9) for collisions with hydrogen. The production rate of ^{26}Al from the two reactions in Table 1 ($1.7 \times 10^{38} \text{ s}^{-1}$) is less than that of $^{12}\text{C}^*$ (4.43 MeV) by only a factor of 12 (allowing for energy-spectrum uncertainty). If this production continues for 10^6 yr (the mean lifetime of ^{26}Al), as it surely must since the Orion complex is considerably older than this, the rate may be expressed as $5.4 \times 10^{51} \text{ atoms Myr}^{-1}$, so that this number of atoms of ^{26}Al should exist in steady state.

We do not know the mass of cloud material in which this production is concentrated. The Compton telescope team estimated⁴ the total cloud mass as $1\text{--}2 \times 10^5 M_\odot$ from the strength of molecular CO emission. The total ^{27}Al content per $10^5 M_\odot$ of solar material is 2.6×10^{56} atoms. For $10^5 M_\odot$ of irradiated clouds, therefore, the steady-state ratio is $^{26}\text{Al}/^{27}\text{Al} = 2.1 \times 10^{-5}$. If the cosmic rays are accelerated by the young stars in hot low-density phases between the cloud components, their local overabundance may result from their diffusion outward, preferentially irradiating certain portions of the clouds. The relatively short range of cosmic rays of 10 MeV per nucleon in the higher magnetic fields within the clouds may restrict the ^{26}Al production to a smaller fraction of their mass (the irradiated surface depth), in which case the $^{26}\text{Al}/^{27}\text{Al}$ ratio would be increased by the same factor in that fraction. For example, if only $10^4 M_\odot$ of cloud are absorbing and stopping the cosmic rays, the ratio in that mass would be $26/27 = 2.1 \times 10^{-4}$. This ratio is much larger than the value of 5×10^{-5} found in some meteoritic samples from the early Solar System¹.

Thus it appears that a cosmic-ray-produced $^{26}\text{Al}/^{27}\text{Al}$ ratio of $\sim 2.1 \times 10^{-5}$ may exist routinely in molecular clouds that are currently undergoing star formation. This value is somewhat larger than that (4×10^{-6}) needed to account for the $2 M_\odot$ of ^{26}Al observed in the entire interstellar medium (ISM), roughly half of which is in molecular clouds. But the contribution of cosmic-ray reactions to this mass depends on the fraction of total molecular-cloud mass that has been bombarded over the past 10^6 yr by cosmic rays. Most molecular clouds are not as active as the Orion complex, so may be bombarded less. An explanation involving these nuclear reactions would, moreover, also have to be consistent with the galactic abundances of Li, Be and B, and must not produce overheating of the clouds. If instead the ^{26}Al is produced by nucleosynthesis in type II supernova occurring at the required rate of 3–4 per century, a past history of galactic infall seems to be required¹⁰ to dilute the old ^{27}Al sufficiently to elevate the mean $^{26}\text{Al}/^{27}\text{Al}$ ratio.

Perhaps more interesting than the issue of the ^{26}Al content of the total ISM is the conclusion that isotopic ratios near several times 10^{-5} may routinely exist in significant portions of molecular clouds that are undergoing star formation. Radiogenic excesses of ^{26}Mg from ^{26}Al decay have become very important to studies of Solar System origin. If the Solar System formed in a molecular cloud on the edge of a cosmic-ray-accelerating region similar to Orion, the $^{26}\text{Al}/^{27}\text{Al}$ ratio of 5×10^{-5} found in meteorites of the Solar System¹ may reflect that cosmic-ray production, which could explain why the early Solar System is tenfold richer in ^{26}Al than is the bulk ISM today. Previous attempts to resolve this question have centred on the idea of a neighbouring asymptotic-giant-branch (AGB) star losing its envelope within the solar nebula, thereby injecting ^{26}Al into the solar cloud¹¹ by causing the solar cloud core to form and to quickly collapse. Cosmic-ray-produced ^{26}Al in the early solar nebula would be distributed inhomogeneously, because the Larmor radius is less than the cloud core radius, so the accelerated cosmic ray will follow flux tubes until mirrored. After collapse of the cloud core, perhaps containing little ^{26}Al , matter continuing to fall onto the solar accretion disk would contain ^{26}Al in a ratio near

TABLE 1 Radioactive abundances in $10^5 M_{\odot}$ clouds

Product	Lifetime (Myr)	Cosmic-ray ion	Reaction	σ (mb)	N (s^{-1})	$^A Z / ^A Z$ (Myr $^{-1}$)
$^{12}C^*$ (4.43 MeV)		^{12}C	(H,p')	200	2×10^{39}	NA
^{26}Al	1.0	^{26}Mg	(H,n)	100	6×10^{37}	7.4×10^{-6}
^{26}Al	1.0	^{28}Si	(H,ppn)	50	1.1×10^{38}	1.4×10^{-5}
^{53}Mn	5.3	^{56}Fe	(H, α)	2	4.2×10^{36}	4.6×10^{-6}
^{53}Mn	5.3	^{54}Fe	(H,2p or pn)	40	5.2×10^{36}	5.8×10^{-6}
^{60}Fe	2.2	^{62}Ni	(H,3p)	20	8.8×10^{34}	1.1×10^{-9}
^{60}Fe	2.2	^{64}Ni	(H,p)	50	5.5×10^{34}	0.7×10^{-9}
^{92}Nb	5.2	^{92}Zr	(H,n)	300	1.4×10^{33}	2.0×10^{-5}
^{107}Pd	9.4	^{108}Pd	(H,pn)	200	1.9×10^{32}	5.3×10^{-6}
^{129}I	118	^{130}Te	(H,pn or nn)	300	1.2×10^{33}	1.6×10^{-5}

The ratios $^A Z / ^A Z$ produced per Myr in the irradiated cloud mass are respectively $^{26}Al / ^{27}Al$, $^{53}Mn / ^{55}Mn$, $^{60}Fe / ^{56}Fe$, $^{92}Nb / ^{93}Nb$, $^{107}Pd / ^{108}Pd$ and $^{129}I / ^{127}I$. For ^{146}Sm see text. Abundance $^A Z$ is product of production rate (N) and 10^6 yr, with no correction for decay; abundance $^A Z$ is that contained in $10^5 M_{\odot}$ of solar composition. Note that the ratios therefore differ for differing irradiation times and for differing mass of irradiated cloud. The heavy-cosmic-ray flux used has solar composition except ^{12}C is depleted by factor of three and ^{26}Mg is enriched by factor of two (ref. 6) relative to ^{28}Si . Cross-sections σ (for the nuclear reaction in column 4) are from literature⁷⁻⁹ or estimated from systematics, and are very spectrum dependent. (NA, not applicable).

$^{26}Al / ^{27}Al = 5 \times 10^{-5}$ because of the cosmic irradiation to which the latter had been exposed. This might explain why Al-rich objects containing vastly different concentrations of ^{26}Al coexist side by side in the regolith of meteorite parent bodies^{12,13}.

Cosmic-ray generation of ^{26}Al should cause this isotope to be correlated with lithium isotope variations⁸, as well as with certain other extinct radioactive nuclides in meteorites. I list in Table 1 five other such isotopes.

^{60}Fe : this isotope, with a mean lifetime of 2.2 Myr, has been detected¹⁴ in meteorites at the level of $^{60}Fe / ^{56}Fe$ (at solidification) = 3.9×10^{-9} . If the isotope is produced by cosmic-ray Ni isotopes, the present model would predict a ratio of 1.8×10^{-9} , suggesting good agreement with the observations at first glance. But the time interval between solar collapse and the solidification of the Chervony Kut meteorite¹⁴ may have been so long as 10 Myr, in which case the initial $^{60}Fe / ^{56}Fe$ ratio would have been 4×10^{-7} , a factor of ~ 100 greater than can be expected from cosmic-ray production. This suggests that we may have to fall back on an injection model, such as that involving an AGB star^{11,15}, for ^{60}Fe (and by implication, for ^{26}Al too). The larger value of $^{60}Fe / ^{56}Fe = 1.6 \times 10^{-6}$ inferred from bulk samples¹⁶ is not known to have been live *in situ*, however; it may be a dead remnant from interstellar dust¹⁷. The significance of the ^{60}Fe abundance and of the time required for solidification of the iron meteorites is therefore critical.

^{129}I : The heavy cosmic-ray flux above will produce a ratio of $^{129}I / ^{127}I = 2 \times 10^{-5}$, primarily from the $H(^{130}Te, ^{129}I)n$ reaction⁸. This value is about one-quarter of the mean meteoritic ratio¹⁸ and is comparable to the variations found from one meteorite to another¹⁸. Because the observed value is considerably less than the expected interstellar values from continuous nucleosynthesis, both on average and within molecular clouds^{19,20}, it appears that not all of the ^{129}I can have been produced in this way. Each of the three r-process extinct radioactivity isotopes (^{107}Pd , ^{129}I , ^{244}Pu) is at least ten times less abundant in meteorites than expected in mean molecular clouds²⁰, so there is no extra source needed for their mean meteoritic abundances. But the inhomogeneous infall onto the solar disk discussed for ^{26}Al above may also account for the inter-meteoritic variations in initial $^{129}I / ^{127}I$. The Te/Mg abundance ratio in the low-energy cosmic rays within Orion clouds is crucial to this issue. Based on the first ionization potentials⁶, fractionation in this ratio is not expected if Orion cosmic rays are indeed representative of the accelerated ISM.

^{92}Nb : the $H(^{92}Zr, ^{92}Nb)n$ reaction⁸ establishes a $^{92}Nb / ^{93}Nb$ ratio of 2×10^{-5} in 10^6 yr for an irradiation capable of producing the $^{26}Al / ^{27}Al$ ratio of 2×10^{-5} . This Nb ratio could in fact be up to 36 times greater after a longer irradiation, owing to the 36 Myr lifetime of ^{92}Nb . An initial meteoritic value of $^{92}Nb / ^{93}Nb =$

1.5×10^{-5} in a high-Nb rutile from the Toluca iron meteorite has recently been reported²¹. Thus the cosmic-ray irradiation model can account for the Nb isotope abundances.

^{53}Mn : the $H(^{56}Fe, ^{53}Mn)He$ cross-section⁸ for low-energy spectra is $\sim 2\%$ of the $H(^{26}Mg, ^{26}Al)n$ cross-section, even though they would be comparable at high energy. Again focusing on the low-energy spectrum, the cosmic-ray model predicts a $^{53}Mn / ^{55}Mn$ ratio of 1×10^{-5} . Meteoritic data for this ratio vary. Measurements on Allende²² suggested an initial ratio of 5×10^{-5} , but the isochron established in an angrite meteorite²³ suggests 1.3×10^{-6} , inconsistent with the present model.

^{107}Pd : the cosmic-ray-generated $^{107}Pd / ^{108}Pd$ ratio is predicted to be 5×10^{-6} , about one-quarter of that inferred in meteorites²⁴. This deficiency, although uncertain given the accuracy of the present nuclear data, could easily be made up by a lengthier irradiation. Alternatively, some ^{107}Pd should also survive in molecular clouds²⁰ from mean galactic nucleosynthesis.

^{146}Sm : Table 1 does not include this extinct nuclide because it can be produced only at a level of 10^{-4} times the concentration of ^{144}Sm from the (p,pn) reaction with ^{147}Sm . This concentration is 100 times too small; but it is expected that residual ^{146}Sm will be present from galactic nucleosynthesis²⁰.

Thus considering that the levels of cosmic-ray-produced extinct radioactive nuclides agree with observations at least within an order of magnitude, it seems reasonable to suggest that acceleration of cosmic rays is ubiquitous in star-forming cloud regions. If this process can account for the amounts of extinct radioactive nuclides in the early Solar System, our view of the dynamical conditions required for solar collapse may have to be revised. If, meanwhile, the cosmic-ray model can account for the level of interstellar ^{26}Al today, it would call into question the idea that our Galaxy experiences three or four core-collapse supernovae every century. $\square$

Received 10 November 1993; accepted 17 January 1994.

- Lee, T., Papanastassiou, D. A. & Wasserburg, G. J. *Astrophys. J.* **211**, L107-L110 (1977).
- Share, G. H. et al. *Astrophys. J.* **292**, L61-L65 (1985).
- Diehl, R. et al. in *Compton Gamma Ray Observatory* (eds Friedlander, M., Gehrels, N. & Macomb, D.) 40-46 (Am. Inst. Phys., New York, 1993).
- Bloemen, H. et al. *Astr. Astrophys.* **281**, L5-L8 (1994).
- Ramaty, R., Kozlovsky, B., & Lingenfelter, R. *Astrophys. J. Suppl. Ser.* **40**, 487-522 (1979).
- Mewaldt, R. A. *Rev. Geophys. Space Phys.* **21**, 295-305 (1983).
- Kozlovsky, B., Lingenfelter, R. & Ramaty, R. *Astrophys. J.* **316**, 801-818 (1987).
- Clayton, D. D., Dwek, E. & Woosley, S. E. *Astrophys. J.* **214**, 300-315 (1977).
- Lang, F. L., Wernitz, C. W., Crannell, C. J., Trombka, J. I. & Chang, C. C. *Phys. Rev.* **C35**, 1214-1227 (1987).
- Clayton, D. D., Hartmann, D. H. & Leising, M. D. *Astrophys. J.* **415**, L25-L29 (1993).
- Cameron, A. G. W. in *Protostars and Planets* Vol. 3, (eds Levy, E. & Lunine, J.) 47 (Univ. of Arizona, Tucson, 1993).
- MacPherson, G. J., Wark, D. A. & Armstrong, J. T. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Mathews, M. S.) 746-807 (Univ. of Arizona, Tucson, 1988).
- Virag, A., Zinner, E., Amari, S. & Anders, E. *Geochim. cosmochim. Acta* **55**, 2045-2062 (1991).

14. Shokulyukov, A. & Lugmair, G. *Science* **259**, 1138–1142 (1993).
15. Wasserburg, G. J., Busso, M., Gallino, R. & Raiteri, C. M. *Astrophys. J.* (in the press).
16. Birc, J.-L. & Lugmair, G. *Earth planet. Sci. Lett.* **90**, 131–135 (1988).
17. Clayton, D. D. *Nature* **257**, 36–37 (1975).
18. Podosek, F. A. & Swindle, T. D. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Mathews, M. S.) 1093–1146 (Univ. of Arizona, Tucson, 1988).
19. Clayton, D. D. *Astrophys. J.* **266**, 381–384 (1983).
20. Clayton, D. D. & Leising, M. D. in *Compton Gamma Ray Observatory Symposium 2* (ed. Gehrels, N.) (Am. Inst. Phys., New York, in the press).
21. Harper, C. L. & Wiesmann, H. *Astrophys. J.* (submitted).
22. Birc, J. L. & Allegre, C. J. *Geophys. Res. Lett.* **12**, 745–748 (1985).
23. Lugmair, G. W., MacIsaac, C. & Shukolyukov, A. *Lunar planet. Sci.* **23**, 823–824 (1992).
24. Kelly, W. R. & Wasserburg, D. J. *Geophys. Res. Lett.* **5**, 1079–1082 (1978).

ACKNOWLEDGEMENTS. We thank A. G. W. Cameron and E. Zinner for their advice. This work was supported by NASA and by the W. M. Keck Foundation.

The curvature elastic-energy function of the lipid–water cubic mesophase

Hesson Chung & Martin Caffrey

Department of Chemistry, The Ohio State University, Columbus, Ohio 43210, USA

CELL and lipid membranes are able to bend, as manifested during membrane fusion and the formation of non-lamellar lyotropic mesophases in water. But there is an energy cost to bending of lipid layers, called the curvature elastic energy. Although the functional form of this energy is known¹, a complete quantitative knowledge of the curvature elastic energy, which is central to predicting the relative stability of the large number of phases that lipid membranes can adopt, has been lacking. Here we use X-ray synchrotron diffraction measurements of the variation of lattice parameter with pressure and temperature for the periodic *la3d* (*Q*²³⁰) cubic phase of hydrated monoolein to calculate the complete curvature elastic-energy function for the lipid cubic mesophase. This allows us to predict the stabilities of different cubic and lamellar phases for this system as a function of composition.

Lipids, when dispersed in water, can display a variety of liquid crystalline phases (Fig. 1). Among these, the inverted cubic and hexagonal phases show a curved interface at the lipid–water boundary. Accessing these highly curved phases from the planar lamellar phase requires a bending of the interface which has an associated curvature energy cost. The same holds when cellular and/or lipid membranes fuse. The early stages of this process require adjacent membranes to join, and in the vicinity of the fusion pore the polar–apolar interface is highly curved.

The curvature elastic-energy is believed to be a critical term governing the stability of non-lamellar phases and the ability of lipid and cellular membranes to bend^{1,2}. To calculate the curvature energy in the lyotropic mesophase, the corresponding monolayer radius of curvature should be defined at the so-called neutral area surface^{2,3}. The latter, representing a surface which does not register a molecular cross-sectional area change upon bending, has been identified in both the inverted hexagonal (*H_{II}*) (ref. 4) and the *la3d* cubic phase⁵. The curvature elastic energy, *E*, can be written in terms of the mean and gaussian curvature at the neutral surface of the lipid monolayer as follows:

$$E = 2k_c \langle H_n^2 \rangle - 2H_0 \langle H_n \rangle + H_0^2 + k_g \langle K_n \rangle \quad (1)$$

where *k_c* is the rigidity constant, *k_g* is the gaussian curvature constant, *H₀* is the spontaneous curvature and *H_n²*, *H_n* and *K_n* represent, respectively, the second moment of mean curvature, mean curvature and gaussian curvature⁶ each evaluated at the neutral surface and $\langle \dots \rangle$ denotes the average obtained by integrating over one of the two monolayers in the unit cell. Although several of the material constants in equation (1) have been determined previously in a variety of lipid systems⁷, neither of the

two curvature constants, *k_c* or *k_g*, has been evaluated in the cubic phase. In what follows, we describe a strategy for, and the results of, such a determination. The curvature elastic-energy function for the *H_{II}* phase, which displays a comparatively simple cylindrical geometry, has been determined⁴. The strategy implemented with the cubic phase is analogous to that used on the *H_{II}* phase. The theory and analysis that follows, were, however, developed and implemented specifically for the cubic phase. It is necessary to consider the cubic phases in the context of minimal surfaces^{6,8}.

The bicontinuous cubic phases are arranged in three-dimensional space such that the bilayer midplane forms an infinite periodic minimal surface (IPMS) (Fig. 1)^{8–11}. Each point on such a surface is characterized by two fundamental curvatures: the mean curvature, *H*, and the gaussian curvature, *K*. The integral of *K* over a closed surface is a topological constant and is related to the genus of the surface as follows¹¹:

$$\int_A K \, dA = 2\pi\chi \quad (2)$$

where *A* and χ are the area and the Euler characteristic of the surface, respectively. For a given IPMS type, the χ value is known⁶. The area of the minimal surface in a cubic unit cell is *A₀a²* where *A₀* is the dimensionless quantity that describes the ratio of the minimal surface area per unit cell to (unit cell volume)^{2/3}, and *a* is the lattice parameter⁶.

As noted above, the midplane of the bilayer forms the IPMS of the bicontinuous cubic phases. However, the neutral area surface is displaced from the minimal surface by a distance, *l_n* (refs 4, 5). The neutral surface in the *la3d* cubic phase of the

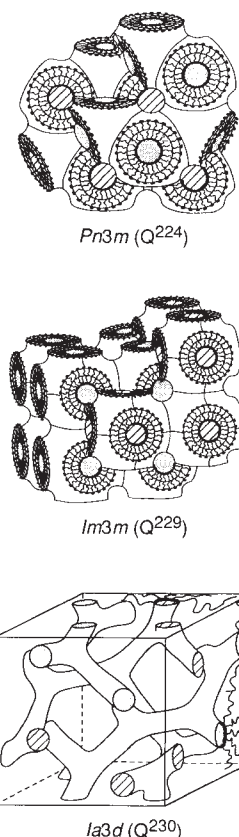


FIG. 1 A schematic representation of the structure of the bicontinuous *Pn3m*, *Im3m* and *la3d* cubic phases (adapted with permission from refs 21 and 22). The surface that sits at the lipid bilayer midplane, separating the two interpenetrating but non-connected water-filled channel networks (shaded and cross-hatched), represents the infinite periodic minimal surface.

FIG. 2 Dependence of hydrated monoolein cubic phase lattice parameter (unit cell length a) on aqueous polyethylene glycol (PEG) concentration, A, and weight fraction (F_1) of monoolein, B, at 25 °C. Weight and volume fraction are used interchangeably as the density of water and of lipid is assumed to be the same. The line of best fit drawn in B is $F_1 = 1.324 - 0.0044669 a$ (correlation coefficient 0.987). The data presented in A and B are combined to provide the dependence of osmotic pressure, P , on the volume of water per lipid molecule, V_w , as shown in C. To this end, the following relation was used to convert PEG concentration in grams of polymer per gram of water, G , to P in dyne cm^{-2} at temperature, T , in the range 5 °C to 40 °C (refs 16, 23):

$$P = -1.31 \times 10^6 T G^2 + 141.8 \times 10^6 G^2 + 4.05 \times 10^6 G \quad (4)$$

V_w ($\text{\AA}^3$) is related to F_1 as follows:

$$V_w = 30(M_l F_w / M_w F_l) \quad (5)$$

where M is molecular weight, F is volume fraction and the subscripts l and w refer to lipid and water, respectively.

METHODS. Diffraction measurements were made in the A1 station at

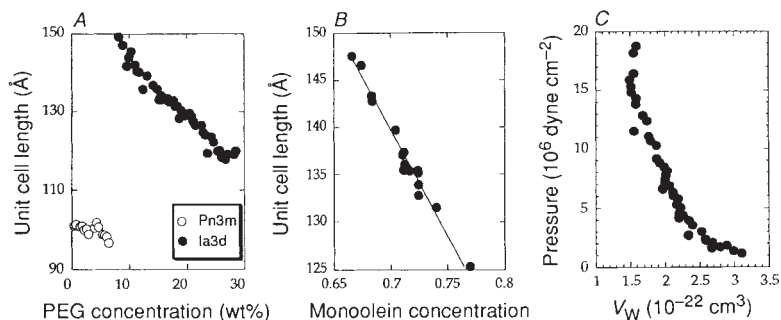
monoolein/water system is located at $8.8 (\pm 1.0)$ and $9.3 (\pm 1.0)$ Å from the terminal methyl group of the hydrocarbon chain at 25 °C and 35 °C, respectively³.

To measure k_c and k_g in a single mesophase, it is necessary to do so in a phase with non-zero gaussian curvature. We have therefore chosen to work with the *Ia3d* (Q^{230} ; Fig. 1) cubic phase of hydrated monoolein. The temperature-composition phase diagram for the monoolein/water system is well established and shows the *Ia3d* cubic phase over a wide temperature and composition range¹⁰. Structural features of this system have been reported^{8,12-15}.

Our first objective was to evaluate $\langle H_n^2 \rangle$, $\langle H_n \rangle$ and $\langle K_n \rangle$ in the *Ia3d* phase of monoolein as E is varied through osmotic stress^{4,16} applied at 25 °C and 35 °C. These average parameters can be calculated by knowing a , the weight fraction of monoolein F_1 , by mesophase mensuration and by using the relations established previously^{6,17}. Overall free energy was adjusted in a controlled way by equilibrating the lipid with aqueous polymer solutions of known concentration and osmotic pressure^{4,16}. In this analysis, we assume that curvature energy is the principal contributor to the overall energy and have neglected other energy terms such as hydration and van der Waals energies. The assumption is based on earlier curvature elastic-energy calculations for the non-lamellar phases^{4,18}. Accordingly, the curvature energy is obtained by numerically integrating the osmotic pressure P , exerted by the polymer solution on the lipid mesophase over the volume of water per lipid molecule V_w , as follows:

$$E = - \int P dV_w = E_{\text{int}} + E_0 \quad (3)$$

where E_{int} is the integrated energy and E_0 is the integration constant. V_w is calculated by knowing F_1 , the molecular weight of the lipid and water and the molecular volume of water ($30 \text{ \AA}^3$ assumed; Fig. 2 legend). The experiments were performed by measuring the dependence of a , (obtained by X-ray diffraction), on P , by adjusting polymer concentration, and on F_1 , and thus V_w , using samples prepared gravimetrically (Fig. 2). This set of measurements combined provided E_{int} . E_0 was set to zero at the excess water phase boundary of the *Pn3m* (Q^{224}) cubic phase of hydrated monoolein (corresponding to F_1 values of 0.606 and 0.625 at 25 °C and 35 °C, respectively; see Fig. 2 of ref. 10). This assumes that the *Pn3m* phase is the most stable state from the point of view of curvature energy under these conditions. In support of this statement, we have established that the cubic phase of monoolein in combination with up to 20 wt% *n*-eicosane persists in excess water at 25 °C (data not shown). Knowing



the Cornell High Energy Synchrotron Source (CHESS) as described²⁴ and by using a Rigaku RU300 rotating copper anode generator (1.54 Å, 54 kV, 240 mA). Monoolein (Nu Chek Prep, Inc., Elysian, Minnesota) had a measured purity of $\geq 99\%$ and was used as received. Aqueous solutions of PEG (Sigma Chemical Co., St. Louis, Missouri) were equilibrated for 3 weeks at room temperature with lipid in the PEG/lipid volume ratio of 10/1. Samples were incubated at 25 °C for >2 h before making the diffraction measurements. Hydrated monoolein samples were prepared as described⁵.

the dependence of E_{int} , $\langle H_n^2 \rangle$, $\langle H_n \rangle$ and $\langle K_n \rangle$ on F_1 (Fig. 3), the constants, H_0 , k_c and k_g , in the curvature energy equation (equation (1)) can be determined by χ^2 fitting. The limiting values of H_0 chosen for use in the fitting protocol were 0 and $0.05 \text{ \AA}^{-1}$, corresponding to spontaneous radius R_0 ($=1/2H_0$) values of ∞ and $10 \text{ \AA}$, respectively. The corresponding values

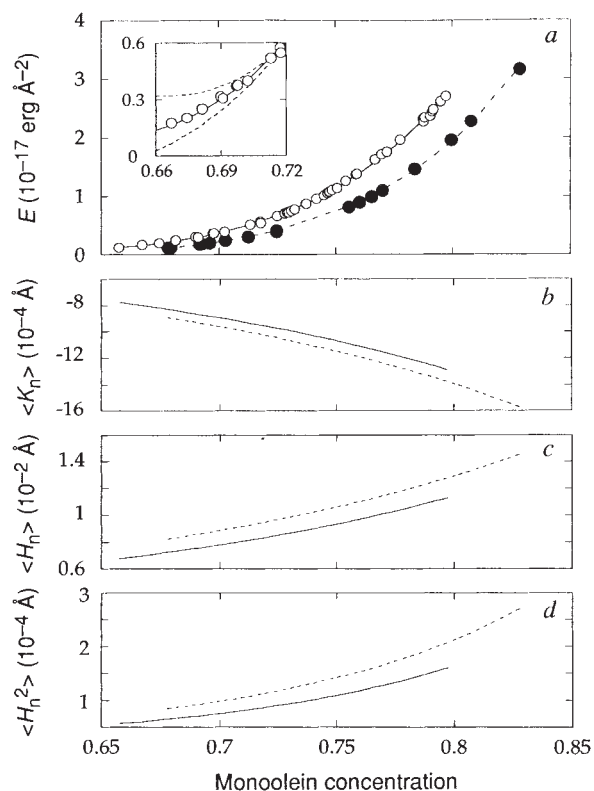


FIG. 3 Monoolein concentration (weight fraction) dependence of a , the curvature energy; b , gaussian curvature; c , mean curvature; and d , second moment of mean curvature of the *Ia3d* phase in hydrated monoolein at 25 °C ($\circ$, —) and 35 °C ($\bullet$, ---). The solid and dashed lines in a are best fitting lines at 25 °C and 35 °C, respectively, obtained by using equation (1) and the constants k_c , k_g and H_0 reported in the text. The inset in a shows the fitted curves when H_0 is fixed at $8.0 \times 10^{-3} \text{ \AA}^{-1}$ (upper dashed line) and k_c is fixed at $3.1 \times 10^{-13} \text{ erg}$ (lower dashed line) along with the experimental data ($\circ$) and best fitting (solid line) at 25 °C.

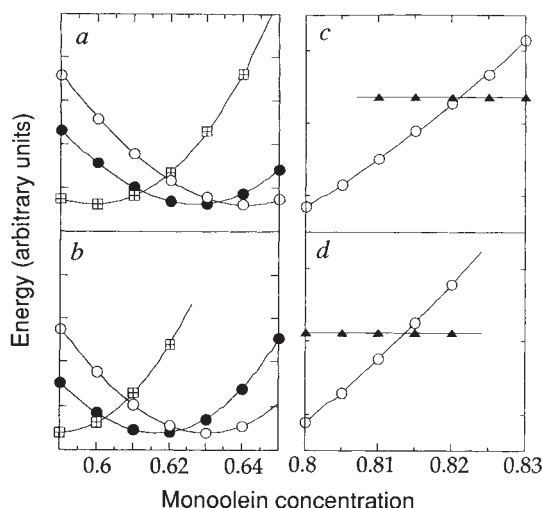


FIG. 4 The energy-composition (weight fraction) diagram of the *Ia3d* (○), *Im3m* (◻) and *Pn3m* (●) cubic phases and the lamellar liquid crystalline (▲) phase of hydrated monoolein at 35 °C (a, c) and 25 °C (b, d). The elastic energy of the lamellar phase is given by $2k_c H_0^2$ at each temperature.

so determined at 25 °C and 35 °C, respectively, are $H_0 = 6.77 (\pm 0.1) \times 10^{-3} \text{ Å}^{-1}$ and $8.01 (\pm 0.1) \times 10^{-3} \text{ Å}^{-1}$ ($R_0 = 73.9 (\pm 2)$ and $62.5 (\pm 2) \text{ Å}$), $k_c = 3.6 (\pm 0.4) \times 10^{-13} \text{ erg}$ and $2.1 (\pm 0.3) \times 10^{-13} \text{ erg}$ and $k_g = 8.8 (\pm 1.8) \times 10^{-15} \text{ erg}$ and $6.6 (\pm 1.8) \times 10^{-15} \text{ erg}$. Accordingly, the curvature constant ratio, k_g/k_c , is $0.024 (\pm 0.008)$ and $0.032 (\pm 0.013)$ at 25 °C and 35 °C, respectively, which compares favourably with the corresponding value of 0.048 reported for the *Im3m* (Q^{229}) phase in a hydrated phospholipid/fatty acid mixture at ~60 °C (ref. 3).

A consequence of knowing the complete curvature elastic-energy function in hydrated monoolein is the ability to use the relation to map out the energy-composition diagram for the various phases as shown in Fig. 4. With increasing hydration, the calculated energy-composition diagram gives the phase sequence L_α -*Ia3d*-*Pn3m*-*Im3m* in perfect agreement with the literature^{12,17,19}. (Ref. 12 should be consulted for a discussion of *Im3m* phase stability.) Furthermore, the *Pn3m*/*Ia3d* and *Ia3d*/ L_α phase boundaries located, respectively, at F_1 values of 0.62 and 0.81 at 25 °C, and at F_1 values of 0.63 and 0.82 at 35 °C in Fig. 4 are in remarkably good agreement with the experimental values of 0.65 and 0.79 at 25 °C, and 0.65 and 0.82 at 35 °C (ref. 10).

Our results indicate that the Helfrich function¹ can account for relative mesophase stability to a first-order approximation. Modifications of the original function to include additional curvature terms and other theoretical approaches²⁰ are being tested currently. We are also evaluating contributions, such as hydration and van der Waals energies, to the overall energy. Using the approach we have introduced here, hypotheses regarding the membrane stabilizing/destabilizing effects of proteins, sterols, antiviral peptides and membrane fusion proteins, to cite just a few examples, can now be tested directly. □

Received 4 November; accepted 30 December 1993.

- Helfrich, W. Z. *Naturf.* **286**, 693–703 (1973).
- Gruner, S. M. *J. phys. Chem.* **93**, 7562–7570 (1989).
- Templer, R. H., Seddon, J. M. & Warrender, N. A. *Biophysical Chem.* (in the press).
- Rand, P. R., Fuller, N., Gruner, S. M. & Parsegian, A. V. *Biochemistry* **29**, 76–87 (1990).
- Chung, H. & Caffrey, M. *Biophys. J.* (in the press).
- Anderson, D. M., Gruner, S. M. & Leibler, S. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5364–5368 (1988).
- Nanavati, C., Markin, V. S., Oberhauser, A. F. & Fernandez, J. M. *Biophys. J.* **63**, 1118–1132 (1992).

- Longley, W. & McIntosh, T. J. *Nature* **303**, 612–614 (1983).
- Scriven, L. E. *Nature* **263**, 123–125 (1976).
- Hyde, S. T., Andersson, S., Ericsson, B. & Larsson, K. Z. *Kristallogr.* **188**, 213–219 (1984).
- Andersson, S., Hyde, S. T., Larsson, K. & Linden, S. *Chem. Rev.* **88**, 221–242 (1988).
- Caffrey, M. *Biochemistry* **26**, 6349–6363 (1987).
- Lindblom, G., Larsson, K., Johansson, L., Fontell, K. & Forsén, S. *J. Am. chem. Soc.* **101**, 5465–5470 (1979).
- Larsson, K. *Nature* **304**, 664 (1983).
- Mariani, P., Luzzati, V. & Delacroix, H. *J. molec. Biol.* **204**, 165–189 (1988).
- Parsegian, V. A., Rand, R. P. & Fuller, N. *Meth. Enzym.* **127**, 400–416 (1979).
- Turner, D. C., Wang, Z.-G., Gruner, S. M., Mannock, D. A. & McElhaney, R. N. *J. Phys. II France* **2**, 2039–2063 (1992).
- Kirk, G. L., Gruner, S. M. & Stein, D. L. *Biochemistry* **23**, 1093–1102 (1984).
- Lindblom, G. & Rilfors, L. *Biochim. biophys. Acta* **988**, 221–256 (1989).
- Kozlov, M. M. & Winterhalter, M. *J. Phys. II France* **1**, 1077–1100 (1991).
- Tate, M. W., Eikenberry, E. F., Turner, D. C., Shyamsunder, E. & Gruner, S. M. *Chem. Phys. Lipids* **57**, 147–164 (1991).
- Seddon, J. M. *Biochim. biophys. Acta* **1031**, 1–69 (1990).
- Michel, B. E. & Kaufmann, M. R. *Pl. Physiol.* **51**, 914–916 (1973).
- Chung, H. & Caffrey, M. *Biophys. J.* **63**, 438–447 (1992).

ACKNOWLEDGEMENTS. We thank D. Anderson, S. Hyde, A. Parsegian, J. Seddon, D. Siegel and R. Templer for evaluating and commenting on this work. An industrial fellowship (H.C.) from The Procter and Gamble Company is acknowledged. This work was supported by the US NIH and the US NSF.

Direct imaging of reptation for semiflexible actin filaments

J. Käs*, H. Strey† & E. Sackmann†

* Division of Experimental Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA

† Technische Universität München, Physik Department, Biophysics Group, 85748 Garching, Germany

ACCORDING to the reptation model of polymer diffusion¹, a polymer chain exhibits snake-like motion through the entangled mesh of surrounding molecules, in which the undulations of the chain are restricted to a tubelike region². The reptation model can account for many of the dynamic properties of entangled polymer solutions and melts, and has received support from observations of block copolymer diffusion across an interface³; but reptative motion has not previously been imaged directly^{4,5}. Here we report such a direct observation of reptation, obtained by video microscopy of fluorescently labelled single, semiflexible filaments of actin in a solution of unlabelled actin filaments. From the restricted thermal undulations of these filaments we can measure the diameter of the confining tube, and we also observe the characteristic thermally excited sliding of the filament out of the end of the tube. We find that the chain self-diffusion coefficient decreases approximately linearly as the filament length increases, in agreement with the reptation model.

Polymerized actin (F-actin) plays an essential role in cell mechanics and cell motility, and is an attractive model for studying the fundamental physical properties of semiflexible polymers⁶. Monomeric actin (G-actin, relative molecular mass $M_r = 42,000$) polymerizes in physiological salt solutions (pH 7.5, 2 mM MgCl_2 , 150 mM KCl) to double-stranded filaments⁷ (F-actin). The F-actin solutions in our experiments exhibit a polydisperse length distribution of ~4–70 μm with a mean length of ~22 μm (ref. 8).

To demonstrate the restriction of the chain dynamics within an entangled polymer solution, we compare in Fig. 1 the transient contours of a free actin filament (Fig. 1a) and an actin filament embedded in a semidilute F-actin solution (Fig. 1b). As first described by Honda *et al.*⁹, a decrease in the amplitudes of the thermally excited undulations is clearly visible for the embedded filament. The restricted chain motions can be understood in terms of the undulations of a filament in a tube formed by the surrounding entangled filaments². As indicated in Fig. 1c, the undulations probe the tube diameter, which allows one to determine its local diameter by measuring the maximum flicker amplitudes.

As demonstrated in Fig. 2, the tube can be imaged and its width determined (as a function of the position along the contour) in a simple way by superimposition of a sufficient number (~ 60) of contours, and determination of the maximum deflection along the contour. The mean diameter, $\langle a \rangle$, of the tube has been determined as a function of the actin concentration c_A . For a semiflexible polymer in an entangled mesh, the apparent tube diameter $\langle a \rangle$ is related to mesh size ξ by $\langle a \rangle \propto \xi (\xi/L_p)^{1/5}$, where L_p is the persistence length¹⁰ and the mesh size ξ describes the mean distance between adjoining polymers. We expect the mesh size to follow approximately the scaling law for stiff rods^{11,12}, $\xi \propto c_A^{-1/2}$ so that $\langle a \rangle \propto c_A^{-3/5}$. Plotting $\langle a \rangle$ as a function of c_A (data not shown) we find $\langle a \rangle \propto c_A^{-0.5 \pm 0.15}$, consistent with the prediction and previous permeation measurements¹³.

As shown in Fig. 3Aa, time sequences of fluorescence micrographs clearly demonstrate the snake-like self-diffusion of the filaments in the F-actin solution. At first the filament slides partially out of the tube at the right side. Then it retracts back into the tube. By this fingering out of the tube the filament explores

new tube segments. Over a longer interval of time (~ 10 min), shown in the example in Fig. 3B, substantial portions of the filament have left the original tube at both ends by repeatedly sliding backwards and forwards. The central section of the molecule, however, remains very close to its original configuration. Thus the reptation movement predicted by de Gennes¹ can be directly imaged for a semiflexible polymer. The central part of the tube is not entirely static, however. Fig. 3B shows a local deviation in the tube contour, indicating the release of a topological constraint when one of the surrounding filaments reptates away. Such events are infrequent, as is expected when the entangling filaments have a size similar to the labelled filaments. This result also illustrates that F-actin solutions form only entangled polymer solutions but no cross-linked networks¹⁴.

The reptation diffusion coefficient along the tube, $D_{||}$, was determined by evaluation of random fingering motion of the chain ends (compare Fig. 3A, b, c). The chain end positions (x_i , y_i), with respect to a local coordinate system with y -axes parallel to the tube axes at the ends, were recorded at time intervals $\Delta t =$

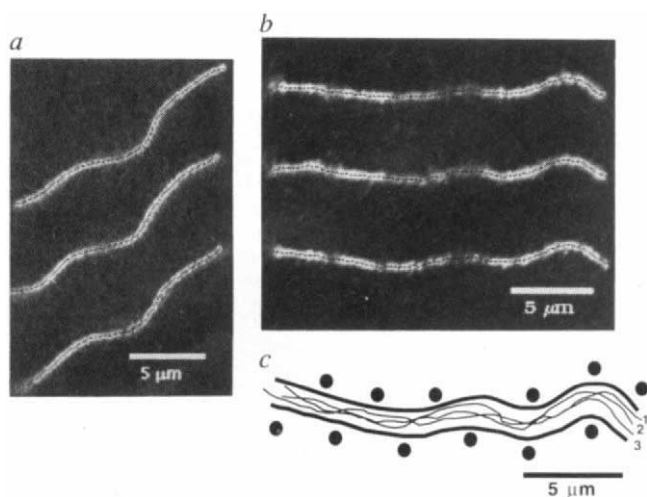


FIG. 1 a, b, Time sequences of an undulating rhodamine-phalloidin-labelled actin filament in a free state (a) and embedded in a non-labelled solution of actin filaments (b, where the actin monomer concentration of the solution $c_A = 0.1 \text{ mg ml}^{-1}$) taken at time intervals of 0.1 s. The dashed black lines in the micrographs show the contour lines determined by image processing. c, Superimposition of transient contour lines (numbered 1, 2, 3) of the filament shown in b. The thick black lines represent the maximum possible undulation amplitudes, which are restricted by the surrounding chains (shown schematically by black circles).

METHODS. Monomeric actin was prepared as previously described¹⁷ and polymerized in the presence of F-buffer (2 mM Tris-HCl, pH 7.4, 1 mM ATP, 0.2 mM CaCl_2 , 1 M KCl, 2 mM MgCl_2) for 12 h. For the sample preparation a very dilute solution (monomer concentration $c_A = 1 \times 10^{-4} \text{ mg ml}^{-1}$) of rhodamine-phalloidin-labelled polymer^{7,18,19} (molar ratio actin/phalloidin 1:1) was mixed with semidilute solutions of non-labelled actin ($c_A = 0.1\text{--}1.4 \text{ mg ml}^{-1}$) by sucking both solutions repeatedly into a pipette. To prevent bleaching of the fluorescent dye, $8 \mu\text{g ml}^{-1}$ catalase, 0.2 mg ml^{-1} glucose, $40 \mu\text{g ml}^{-1}$ glucose oxidase and 0.5 vol% mercaptoethanol were added to the sample and the sample was degassed for 30 min. The selective labelling of filaments in an unlabelled F-actin solution allows observation of the dynamics of a single polymer. Video microscopy and detection of transient contours ($x(s)$, $y(s)$) (s , arc length) of the fluorescence-labelled filaments by image processing were performed as previously described¹⁷. Only filaments which were in the focal plane of the microscope were analysed. The positions ($x(s)$, $y(s)$) along the contour could be located with an accuracy of $\pm 50 \text{ nm}$ and was limited only by the signal-to-noise ratio.

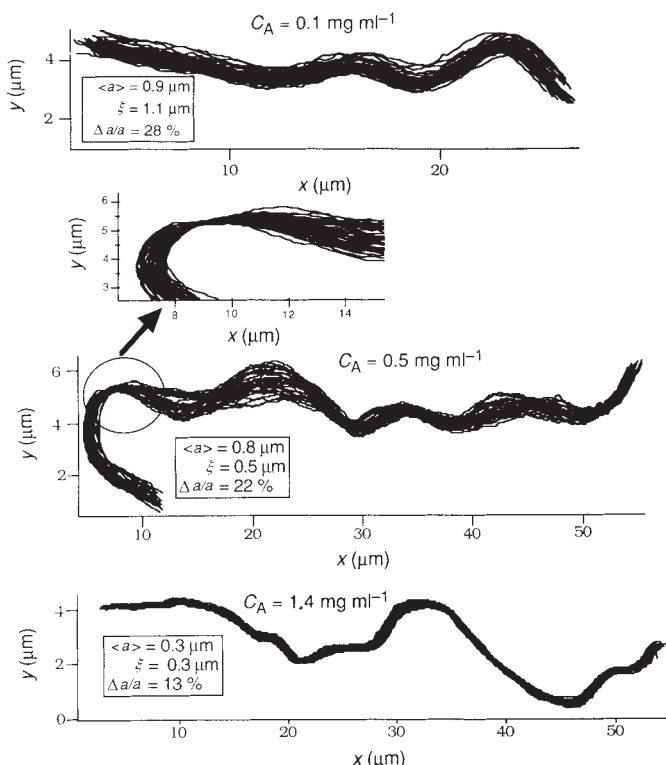


FIG. 2 Determination of the local diameter of tube formed around a labelled actin filament by surrounding filaments. The tubes were imaged by the superimposition of 60 traces of the contour taken at 0.1-s intervals. The mean tube diameter $\langle a \rangle$ is determined by averaging the width of the dark bands along the contour. Above each tube the actin monomer concentration c_A of the surrounding F-actin solution is noted. In the boxes are given values of the average tube diameter $\langle a \rangle$, together with the mesh size ξ determined by permeation studies¹³. The fluctuations of the tube diameter $\Delta a / \langle a \rangle$ are also presented. As shown in the enlargement, the tube diameter exhibits local knots demonstrating large fluctuations $\Delta a / \langle a \rangle$ of the mesh size. The existence of knots was already predicted by Wendel and Noolandi in their modified reptation model²⁰. The large variance of the mesh size is typical for semidilute solutions, because of the large concentration fluctuations in this regime¹⁵. The data of the inserts also show that the fluctuation of the tube diameter, $\Delta a / \langle a \rangle$, is smaller for the highest concentration.

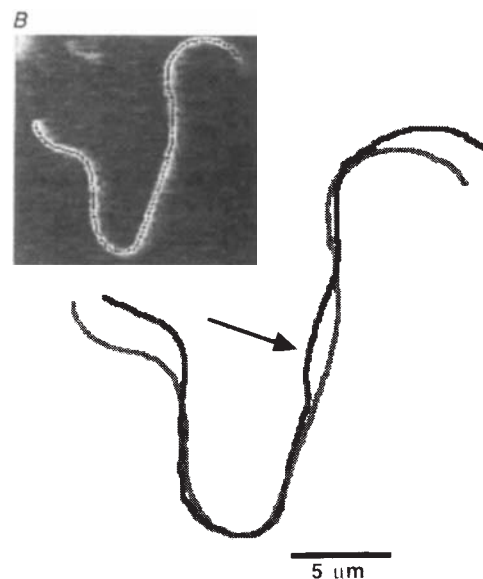
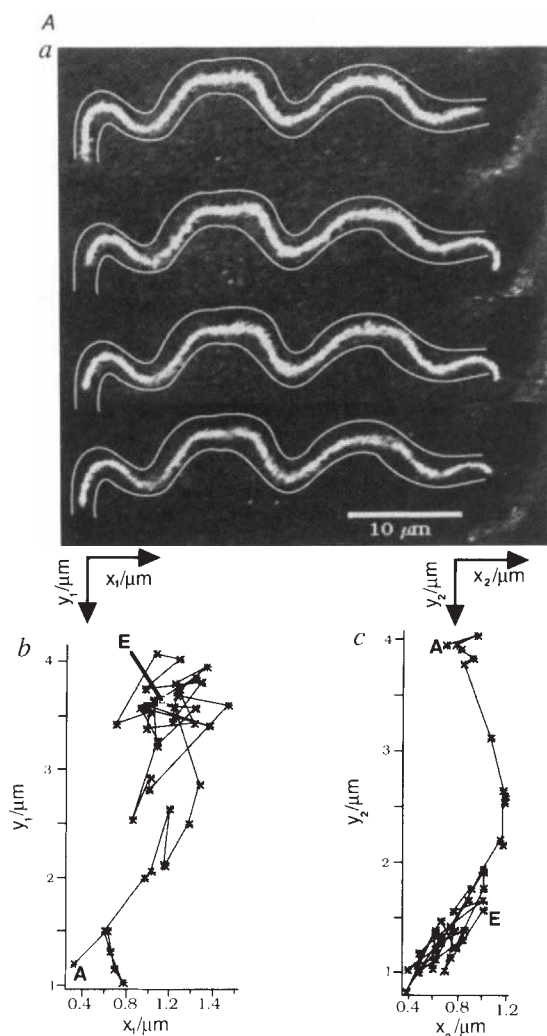


FIG. 3 Aa, Time sequence of the reptation motion of a semiflexible actin filament. Snapshots of the filament embedded in a solution of actin filaments of monomer concentration $c_A = 0.1 \text{ mg ml}^{-1}$ are taken at time intervals of 20 s. The schematic diameter of the tube drawn to guide the eye (thin white lines) is three times larger than the real mean diameter ($\langle a \rangle = 0.9 \mu\text{m}$). Ab, Ac, Random walk of chain ends (x_i, y_i) ($i=1, 2$) of the filament of contour length $L_C = 51 \mu\text{m}$ shown in a. The y_i denote the local axes of the tube at the two ends. Positions (x_i, y_i) of end-points were recorded at time intervals of $\Delta t = 0.84 \text{ s}$. The time sequences started at points A and ended at points E. Note that single jumps of the two ends must not be correlated, because they are driven by the thermal undulations, but on long timescales motion of the two ends is correlated and proves that the filament is really sliding out of the tube. B, Traces of the contour of a tube surrounding the filament shown in the photograph. The first trace (grey) is taken at the beginning of the experiment. The second trace (black) is taken after 7.5 min. It may be seen that the filament has left the original tube at both ends, which happens by the filament repeatedly sliding backward and forward. The central part of the tube is not entirely static. The arrow shows a local deviation in the tube contour, indicating release of constraint when one of the surrounding filaments reptates away.

0.84 s. $D_{\parallel}$ was determined as the arithmetic mean of the diffusion coefficients parallel to the tube at both ends (D_1 and D_2) according to

$$D_{\parallel} = \frac{1}{2} (D_1 + D_2) \\ = \frac{1}{2} \left(\frac{1}{(N-1)} \sum_{i=2}^N \frac{(y_1(i) - y_1(i-1))^2}{2 \times \Delta t} + \frac{1}{(N-1)} \sum_{i=2}^N \frac{(y_2(i) - y_2(i-1))^2}{2 \times \Delta t} \right) \quad (1)$$

where N is the number of steps and was $N \approx 40$. The accuracy of the measurement is $\Delta D = \pm 2 \times 10^{-14} \text{ m}^2 \text{ s}^{-1}$. Our experiments clearly show that semiflexible chains embedded in entangled solutions reptate over small timescales with respect to the reptation time—the time needed for the chain to disengage from the tube—but large with respect to fluctuations inside the tube. To prove that there is reptation movement, we show that the motion of both ends of the filament are correlated.

While a c_A -dependence is not observed within experimental error, $D_{\parallel}$ decreases approximately linearly with increasing contour length L_C as shown in Fig. 4. Because of the lack of a theory for semiflexible polymers, we compared the data with the following formula predicted for stiff rods¹⁵: $D_{\parallel} =$

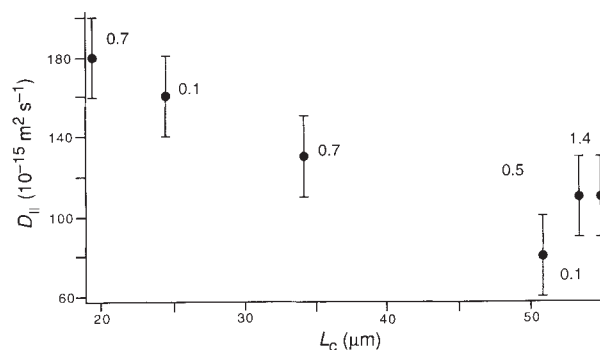


FIG. 4 Measured contour-length-dependency of the reptation diffusion coefficient $D_{\parallel}$. A decrease of the diffusion coefficient $D_{\parallel}$ with increasing contour length L_C is clearly visible. The numbers beside the data points show the actin monomer concentration c_A (in mg ml^{-1}) of the surrounding F-actin solution.

$(k_B T \ln(L_C/b))/(2\pi\eta_s L_C)$, where $b=7$ nm is the diameter of the filament, k_B is the Boltzmann constant, T is the temperature and η_s is the viscosity of the buffer, which could be approximated by the viscosity of water. We found that the theoretical values and the measured data are of the same order of magnitude, and that the experimental values are lower by a factor of ~ 2 .

The flexibility of polymers can be characterized by the persistence length L_p . This is based on the idea that the correlation between the orientations of the local tangent $t(s)$ decays with the distance s along the filament contour according to $\langle t(s)t(s') \rangle = \exp(-|s-s'|/L_p)$. The persistence length L_p of the embedded

filaments was determined from end-to-end distance, R_{end} , measurements using the Kratky-Porod model¹⁵ for semiflexible chains

$$R_{\text{end}}^2 = 4L_p^2[L_C/L_p - 2(1 - \exp\{-L_C/2L_p\})] \quad (2)$$

The factor 4 holds for chains confined to two dimensions¹⁶. End-to-end distances were measured for four different concentrations, evaluating ~ 25 chains for each concentration. By a least-squares fit of the data to equation (2), we find a persistence length of $L_p = 4 \pm 1 \mu\text{m}$ which confirms the semiflexible character of the actin filaments. $\square$

Received 13 October 1993; accepted 26 January 1994.

- de Gennes, P. G. *J. chem. Phys.* **55**, 572–579 (1971).
- Edwards, S. F. *Proc. phys. Soc.* **92**, 9–13 (1967).
- Russell, T. P. et al. *Nature* **365**, 235–237 (1993).
- Higgins, J. & McLeish, T. *Nature* **365**, 205–206 (1993).
- Lodge, T. P., Rotstein, N. A. & Prager, S. *Adv. chem. Phys.* **79**, 1–132 (1990).
- Sackmann, E. *Macromolec. Chem. Phys.* **194**, 7–28 (1994).
- Alberts, B. et al. *Molecular Biology of the Cell* 2nd edn (Garland, New York, 1989).
- Kaufmann, S., Käs, J., Goldmann, W. H., Sackmann, E. & Isenberg, G. *FEBS Lett.* **314**, 203–205 (1992).
- Honda, H., Nagashima, H. & Asakura, S. *J. molec. Biol.* **191**, 131–133 (1986).
- Semenov, A. N. *J. chem. Soc., Faraday Trans. 2* **82**, 317–329 (1986).
- de Gennes, P. G., Pincus, P., Velasco, R. M. & Brochard, F. *J. Phys., Paris* **37**, 1461–1473 (1976).
- Graessley, W. W. *J. polym. Sci.* **13**, 17–34 (1980).
- Schmidt, C., Bärmann, M., Isenberg, G. & Sackmann, E. *Macromolecules* **22**, 3638–3649 (1989).

- Janmey, P. A. et al. *Biochemistry* **27**, 8218–8227 (1988).
- Doi, M. & Edwards, S. F. *The Theory of Polymer Dynamics* 2nd edn (Clarendon, Oxford, 1988).
- Landau, L. D. & Lifshitz, E. M. *Statistical Physics* 3rd edn Part 1 544–547 (Pergamon, Oxford, 1980).
- Käs, J., Strey, H., Bärmann, M. & Sackmann, E. *Europhys. Lett.* **21**, 865–870 (1993).
- Yanagida, T., Nakase, M., Nishiyama, K. & Oosawa, F. *Nature* **307**, 58–60 (1984).
- Kishino, A. & Yanagida, T. *Nature* **334**, 74–76 (1988).
- Wendel, H. & Noolandi, J. *Macromolecules* **15**, 1318–1320 (1982).

ACKNOWLEDGEMENTS. We are grateful for discussions with R. Ezzell, E. Frey, P. Janmey, D. Nelson, J. Noolandi and P. Pincus. We also thank P. Allen, D. Jackson, K. Kremer and T. Maggs for their help. J.K. especially thanks T. Duke and S. Leibler for help and advice. The work was supported by the Deutsche Forschungsgemeinschaft and by the Fonds der chemischen Industrie.

Purification of C_{60} and C_{70} by selective complexation with calixarenes

Jerry L. Atwood*, George A. Koutsantonis† & Colin L. Raston†‡

* Department of Chemistry, University of Alabama, Tuscaloosa, Alabama 35487, USA

† Faculty of Science and Technology, Griffith University, Nathan, Brisbane 4111, Australia

MOLECULAR complexation of fullerenes in host-guest complexes has the potential to afford efficient large-scale purification of fullerenes¹. This method would avoid the losses inherent in the chromatographic techniques currently in use, which arise from irreversible absorption on the stationary phase². Several host-guest interactions have been reported for C_{60} , including the formation of complexes with 1,4-hydroquinone^{3,4}, azacrown compounds⁵, certain water-soluble macrocycles such as γ -cyclodextrin^{6–8}, and complexation between C_{60} and an iridium phosphine in which pendant groups attached to a phosphorus atom form a cradle for an adjacent C_{60} molecule⁹. Other relevant studies are the inclusion of C_{60} into a microporous aluminophosphate¹⁰ and the formation of a charge-transfer complex between C_{60} and a thiafulvalene¹¹. Here we show that both C_{60} and C_{70} will form discrete complexes with calixarenes, bowl-shaped macrocycles with hydrophobic cavities. Complexation of p -Bu'-calix[8]arene with a mixture of the toluene extract of 'crude' fullerene soot, followed by a series of recrystallizations, affords $>99.5\%$ pure C_{60} . Our approach could lead to a substantial reduction in the cost of purifying C_{60} and C_{70} .

Calixarenes are macrocycles with well ordered cyclic arrays. The phenolic-OH groups are H-bonded and directed towards the centre of the cycle (1 and 2 in Fig. 1). The array has some degree of flexibility¹², and may adjust in favourable cases for

docking with fullerenes. Toluene solutions of p -Bu'-calix[8]arene, 1, $R = \text{Bu}'$, form a sparingly soluble brown/yellow precipitate in the presence of a toluene solution of purified C_{60} (ref. 13) (Fig. 2) analysing as the 1 : 1 complex, 3. In the presence of toluene solutions of crude fullerene mixtures a similar precipitate results, which was shown by ultraviolet-visible spectro-

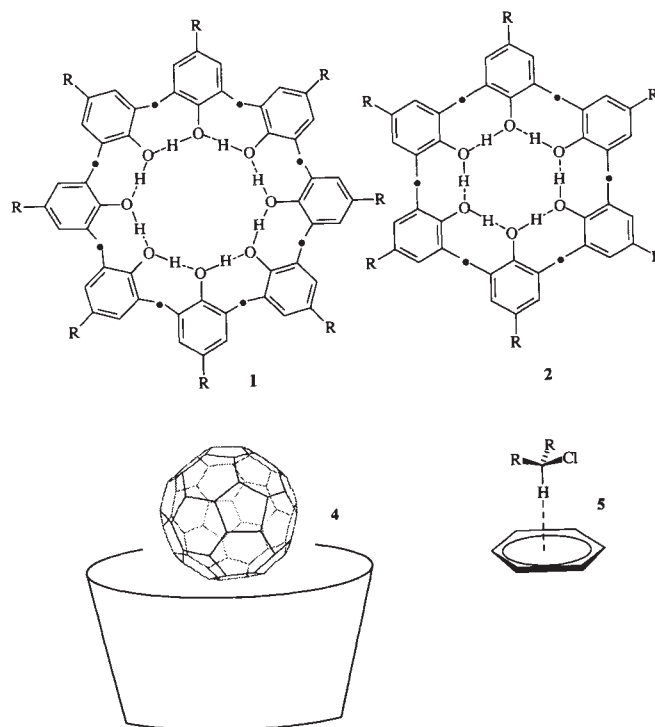


FIG. 1 Structural formulae for p -R-calix[8]arene, 1, and p -R-calix[6]arene, 2; diagrammatic representation of complex 3 as a 'ball and socket' nano-structure, 4; and the interaction of a solvent hydrogen atom with an arene π -cloud, 5.

‡ To whom correspondence should be addressed.

scopy (Fig. 3) and fast-atom-bombardment (FAB) mass spectrometry to be the same complex containing some C_{70} . The fullerene mixture retrieved from this precipitate by addition of chloroform (Fig. 2) was shown by HPLC¹⁴ to be based exclusively on C_{60} and C_{70} , 89% and 11% respectively, which represents ~90% of the C_{60} content of the soot. One recrystallization of the precipitate, from toluene, enriched the C_{60} content to 96% with ~10% of the fullerenes in the mother liquor as 26% C_{60} and 74% C_{70} . A second recrystallization gave >99.5% purity C_{60} , again with ~10% of the fullerenes in the mother liquor, as 72% C_{60} and 28% C_{70} .

The recovery of high-purity C_{60} is essentially quantitative if mother liquor solutions are recycled back to the crude fullerene mixture, along with *p*-Bu^t-calix[8]arene recovered from chloroform degradation of the complex. In addition, *p*-Bu^t-calix[8]arene has a higher affinity for C_{60} than C_{70} ; we note that purified C_{70} (ref. 13) does not readily form a complex with *p*-Bu^t-calix[8]arene. Presumably the cavity of the calixarene¹⁴ is too small for favourable binding of higher fullerenes. The infrared spectrum of complex **3** is a superposition of that for the calixarene and C_{60} , except for the ν_{OH} stretching region (Fig. 4). Nevertheless, the H-bonding network is retained, and the fullerene most likely resides in the cavity, as a 'ball and socket' nano-structure, **4** in Fig. 1.

Complex **3** rapidly decomposes in chloroform; this allows easy

removal of C_{60} as a black precipitate (>95% recovery) from the calixarene which crystallizes as a suspended solid (relative densities: fullerene C_{60} > chloroform > *p*-Bu^t-calix[8]arene). Fullerene C_{60} is only sparingly soluble in chloroform (0.16 mg ml⁻¹)¹⁶, and the procedure allows for recovery of the calixarene, either as a solid directly, or as a mixture of fullerene and calixarene on removal of the solvent *in vacuo*. Dichloromethane and 1,2-dichloroethane also result in decomposition of the complex, whereas the complex is stable in carbon tetrachloride and *t*-butyl chloride. The absence of hydrogen atoms attached to the carbon bearing a chloro group in these latter solvents (unlike that for the foregoing chlorinated solvents in Fig. 1) relates to the ability of such hydrogen atoms to form significant interactions with aromatic π -rings, **5** (refs 17, 18). This could involve either the calixarene, or more likely the fullerene itself, and at the expense of complexation.

In aromatic solvents the choice of toluene for complex formation is critical in our process. Complex **3** is slightly soluble in toluene at ambient temperature, affording yellow solutions; above ~80 °C this complex decomposes, as judged by the formation of magenta solutions characteristic of free C_{60} in toluene, and ultraviolet-visible spectroscopy. Dilute solutions of **3** in toluene at ambient temperature also contain free C_{60} (see Fig. 3). Xylenes and mesitylene rapidly decompose the complex at 20 °C, yielding free fullerene and calixarene on evaporation. Greater solvation of the C_{60} by the more electron-rich arenes may be at the expense of complex formation. Removal of the calixarene from the fullerene can also be achieved by refluxing a toluene solution over sodium hydroxide pellets for ~10 min (see Fig. 2); the calixarene separates as a sodium salt which fails to form a complex on cooling, presumably because of disruption of the cavity.

Calix[6]arene, **2**, R=H, forms a sparingly soluble 1:2 complex with C_{60} in toluene, complex **6** (Fig. 2), but not with C_{70} in the same solvent. Treatment of a toluene solution of crude fullerenes with the same calixarene yields a highly crystalline material containing both C_{60} and C_{70} , 67% and 33% respectively (HPLC analysis of NaOH-treated toluene solutions of the com-

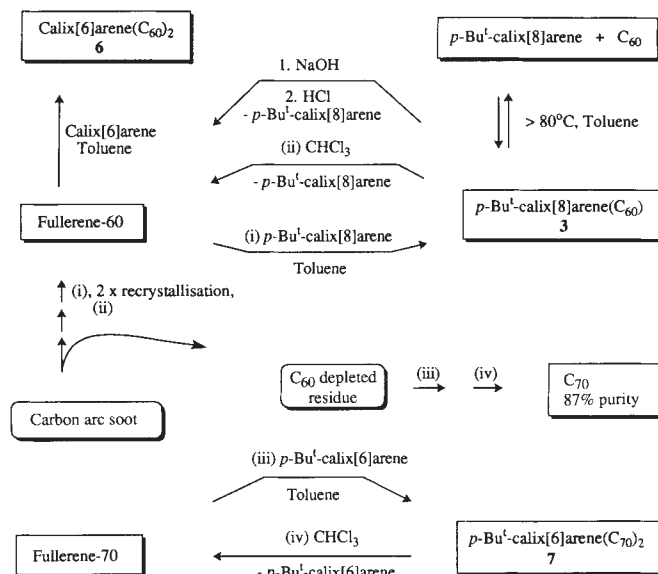


FIG. 2 Reagents and conditions for the complexation of *p*-Bu^t-calix[8]arene, *p*-Bu^t-calix[6]arene and calix[6]arene with C_{60} and C_{70} . METHODS. Synthesis of complex **3** and >99.5% purity C_{60} . In a typical experiment ~7.5 g of fresh carbon-arc soot was stirred with 250 ml of toluene for 1 h then the mixture was filtered, and 1.0 g of the *p*-Bu^t-calix[8]arene added. After refluxing for 10 min the mixture was filtered, seeded with crystals of **3** then allowed to stand at ~20 °C overnight. The yellow/brown plates of the complex were then collected and recrystallized twice from toluene (1.0 g from 80 ml of toluene), 90% yield (analysis: found C 88.20, H 5.25%; calculated C 88.07, H 5.59%). Addition of chloroform (5 ml) to the complex (0.85 g) afforded a precipitate of C_{60} (0.28 g, 92% recovery from the recrystallized complex). Synthesis of Calix[6]arene(C_{60})₂, **6**. To a refluxing solution of C_{60} (5 mg) in toluene (5 ml) was added calix[6]arene (4.4 mg). The hot solution was rapidly filtered and slowly cooled overnight yielding black prisms (5.5 mg, 77% yield, analysis: found C 94.18, H 1.87%; calculated C 93.63, H 1.75%). Synthesis of *p*-Bu^t-calix[6]arene(C_{70})₂, **7**. To a refluxing solution of C_{70} (5 mg) in toluene (2 ml) was added *p*-Bu^t-calix[6]arene (5.8 mg). The hot solution was rapidly filtered and slowly cooled overnight yielding red/brown needles (2.5 mg, 31% yield, analysis: found C 92.48, H 3.19%; calculated C 93.2, H 3.19%).

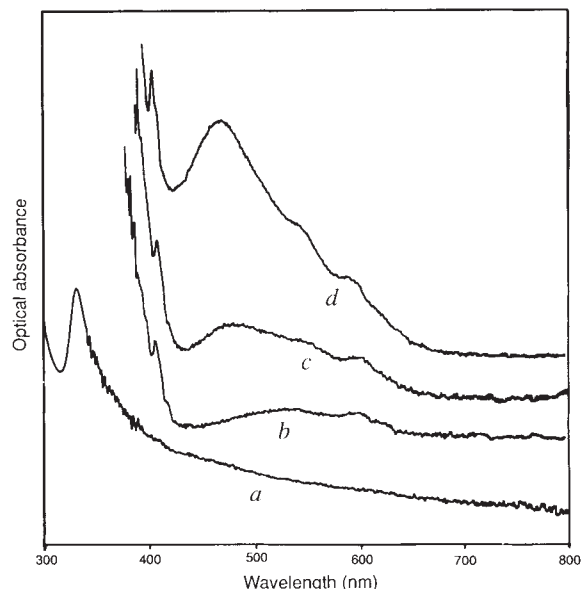


FIG. 3 Ultraviolet visible absorption spectra of a, a carbon tetrachloride solution of complex **3**, b, toluene solution of complex **3**, c, toluene solution of the precipitate formed from crude fullerene mixtures and *p*-Bu^t-calix[8]arene, and d, a dichloromethane solution of fullerenes recovered from the mother liquor from the first recrystallization of the complex in c.

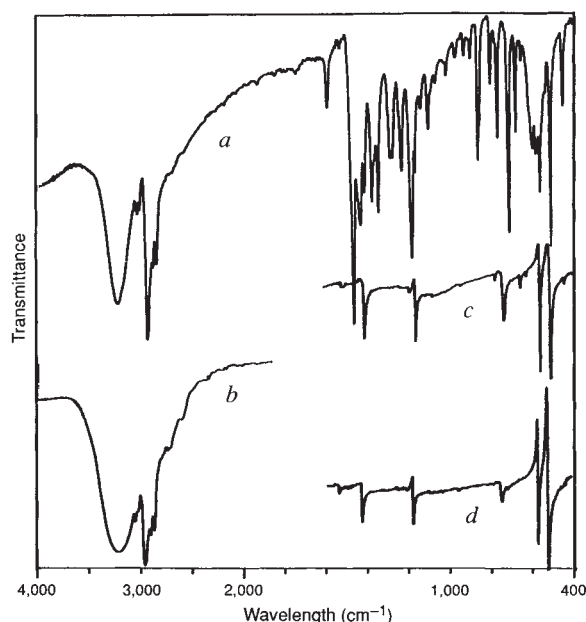


FIG. 4 Infrared spectra (KBr discs) of *a*, complex **3**, *b*, *p*-Bu'-calix[8]arene, excluding the fingerprint region which matches the spectrum of **3** except for the bands corresponding to fullerene C₆₀, *c*, the fullerene mixture derived from decomposition of the precipitate formed between *p*-Bu'-calix[8]arene and a toluene solution of crude fullerene, and *d*, C₆₀ purified by complexation/recrystallization (the band at 760 cm⁻¹ corresponds to chloroform).

plex). Thus there is no discrimination between the two fullerenes. We also find that *p*-Bu'-calix[6]arene, **2**, R = Bu', forms a 1:2 complex with C₇₀, complex **7** (Fig. 2). Treatment of the fullerene residue (depleted of most of the C₆₀ as complex **3**) with *p*-Bu'-calix[6]arene yields a precipitate which, on treatment with chloroform, yields a material rich in C₇₀ (87% along with 13% C₆₀). Thus we are *en route* to purification of this fullerene also by complexation.

Overall our findings demonstrate the ability to selectively recover high-purity C₆₀ from carbon-arc soot by simple and inexpensive complexation/recrystallization procedures, and also the enrichment of C₇₀ content of the residues. Moreover, the formation of complex **3** as a sparingly soluble material is an attractive method of purifying C₆₀ from reaction mixtures, as we have discovered for reaction mixtures derived from treating anthracene with C₆₀ (refs 19, 20). Further studies are in progress on the use of other calixarenes and related molecules to selectively bind higher fullerenes, and ultimately lead to their purification.

We estimate that our approach will lead to a 50-fold reduction in the cost of production of high purity C₆₀, noting the current Aldrich price of \$2,800 per gram, and taking into account (1) the initial capital outlay, including that for recyclable *p*-Bu'-calix[8]arene, \$1.43 per gram, (2) the relatively inexpensive consumables, including toluene and chloroform, although these are largely recoverable, and graphite rods (~\$3.00 per rod, yielding ~1.0 g of C₆₀), and (3) personnel requirements. Although our complexation approach has not yet yielded pure C₇₀, this should be possible in principle; and a similar reduction in the cost of C₇₀ production is likely given our present results and the Aldrich list price of \$10,700 per gram. □

- Ermer, O. *Helv. chim. Acta* **74**, 1339–1351 (1991).
- Ermer, O. & Robke, C. *J. Am. chem. Soc.* **115**, 10077–10082 (1993).
- Diederich, F. et al. *Angew. Chem. int. Edn. engl.* **31**, 1599–1602 (1992).
- Anderson, T., Nilsson, K., Sundahl, M., Westman, G. & Wennerstrom, O. *J. chem. Soc., chem. Commun.* 604–606 (1992).
- Kutner, P., Bouies, P. & Kadish, K. M. *J. electrochem. Soc.* **139C**, 243 (1992).
- Williams, R. M. & Verhoeven, J. W. *Recl Trav. chim. Pays-Bas Belg.* **111**, 531–532 (1992).
- Balch, A. L., Catalona, V. J., Lee, J. W. & Olmstead, M. M. *J. Am. chem. Soc.* **114**, 5455–5457 (1992).
- Anderson, M. A. et al. *J. chem. Soc., chem. Commun.* 533–536 (1993).
- Izuoka, A. et al. *J. chem. Soc., chem. Commun.* 1472–1473 (1992).
- Gutsche, C. D. *Calixarenes* (ed. Stoddart, J. F.) (Monogr. in Supramolecular Chem., Royal Society of Chemistry, Cambridge, 1989).
- Scrivens, W. A. & Tour, J. M. *J. org. Chem.* **57**, 6932–6936 (1992).
- Meier, M. S. & Selegue, J. P. *J. org. Chem.* **57**, 1924–1926 (1992).
- Gutsche, C. D., Gutsche, A. E. & Karaulov, A. I. *J. Incl. Phenom.* **3**, 447–451 (1985).
- Ruoff, R. S., Tse, D. S., Malhotra, R. & Lorents, D. C. *J. phys. Chem.* **97**, 3379–3383 (1993).
- Jorgenson, W. L. & Severence, D. L. *J. Am. chem. Soc.* **112**, 4768–4774 (1990).
- Atwood, J. L., Hamada, F., Robinson, K. D., Orr, G. W. & Vincent, R. L. *Nature* **349**, 683–685 (1991).
- Schluter, J. A. et al. *J. chem. Soc., chem. Commun.* 972–974 (1993).
- Tsuda, M., Ishida, T., Negami, T., Kurono, S. & Ohashi, M. *J. chem. Soc., chem. Commun.* 1296–1298 (1992).

ACKNOWLEDGEMENTS. We thank S. Bottle and M. Walsh for assistance in recording HPLC and UV-visible spectra, and acknowledge financial support from the Australian Research Council and the US National Science Foundation.

Homogeneous catalytic hydrogenation of supercritical carbon dioxide

Philip G. Jessop, Takao Ikariya & Ryoji Noyori*

ERATO Molecular Catalysis Project, Research Development Corporation of Japan, 1247 Yachigusa, Yakusa-cho, Toyota 470-03, Japan

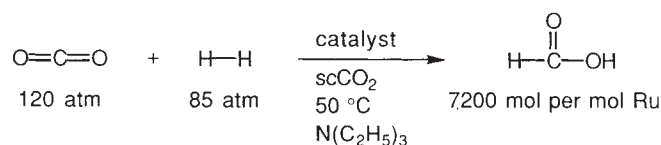
THE use of carbon dioxide as a starting material for the synthesis of organic compounds has long been a goal for synthetic chemists. The hydrogenation of carbon dioxide to formic acid, methanol and other organic substances is particularly attractive, but has remained difficult. This route to formic acid has been described recently, based on the use of organometallic rhodium catalysts in dimethyl sulphoxide¹ and aqueous² solvents. We report here the efficient production of formic acid in a supercritical mixture of carbon dioxide and hydrogen containing a catalytic ruthenium(II) phosphine complex. The use of a supercritical phase, in which hydrogen is highly miscible, leads to a very high initial rate of reaction—up to 1,400 moles of formic acid per mole of catalyst per hour. The same reaction under identical conditions but in liquid organic solvents is much slower. Our results suggest that supercritical fluids represent a promising medium for homogeneous catalysis.

When carbon dioxide is heated beyond its critical point, 31 °C and 73 atm, the gas and liquid phases merge into a single supercritical phase (scCO₂). Supercritical fluids, and CO₂ in particular, are currently used in a wide range of applications, such as chromatography and extraction processes^{3,4}. They represent unusual media for chemical reactions because they combine characteristics associated with gas phase reactions, such as miscibility with other gases, high mixing rates and relatively weak molecular association (in comparison to ordinary condensed phases), resulting in high reactivity of catalysts and reactants, with the advantageous properties of liquid solvents such as their ability to dissolve and transport organic compounds. Many other advantages have been mentioned in the review literature^{5–7}. With scCO₂ as a reaction medium, dispensing with liquid organic solvents is also attractive because it avoids the inflammability, toxicity, and disposal costs associated with the use of such solvents. Furthermore, CO₂ separation from the reaction mixture is

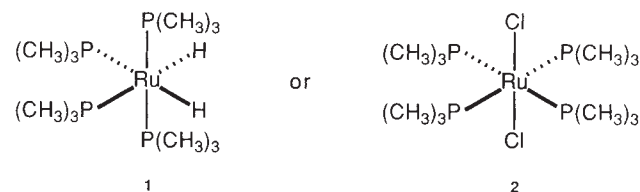
Received 12 November 1993; accepted 7 February 1994.

- Diederich, F. & Rubin, Y. *Angew. Chem. int. Edn. engl.* **31**, 1101–1123 (1992).
- Taylor, R. & Walton, D. R. M. *Nature* **363**, 685–693 (1993).

* To whom correspondence should be addressed at: Department of Chemistry, Nagoya University, Chikusa, Nagoya 464-01, Japan.



catalyst:



energy-efficient and the product is directly obtainable by simple pressure reduction. Despite these obvious benefits, almost no research has been reported on the use of scCO_2 as a medium for homogeneous transition-metal-catalysed reactions (for a notable exception, see ref. 8). The present results demonstrate that these methods can be successfully applied to an important chemical reaction with excellent catalyst performance.

The hydrogenation of carbon dioxide to formic acid has attracted attention as a way of regenerating organic compounds and of storing H_2 . Although the conversion, $\text{CO}_2(\text{g}) + \text{H}_2(\text{g}) \rightleftharpoons \text{HCOOH} (1)$, occurs with $\Delta G^\circ = +33 \text{ kJ mol}^{-1}$ (ref. 9), addition of a base and use of high pressures can shift the equilibrium in favour of product formation. Homogeneous hydrogenation using organometallic molecular catalysts indeed allows hydrogenation of carbon dioxide to formic acid in the presence of basic agents such as tertiary amines. In pioneering investigations, however, the catalytic efficiency remained unsatisfactory¹⁰⁻¹³ for practical purposes, though recent reports showed more promise^{1,2,14} (yielding up to about 3,440 moles of formic acid per mole of catalyst).

We found that certain $\text{Ru}(\text{II})$ -phosphine-complex catalysts are highly active for the hydrogenation of carbon dioxide in a supercritical state (Fig. 1). The catalysts were chosen for their known solubility in hexane, which, like scCO_2 , is nonpolar and weakly solvating. In a stainless steel vessel, $\text{RuH}_2(\text{P}(\text{CH}_3)_3)_4$ (compound 1 in Fig. 1), excess triethylamine, and a trace amount of water (typically 1:3,100:40 mole ratio) were dissolved in a supercritical mixture of H_2 (85 atm) and CO_2 at a total pressure of 200–220 atm and at 50°C . In order to prevent the formation of a liquid triethylamine phase, the amount of the amine was kept at $\leq 1/3$ of the amount of the amine which could be completely dissolved in scCO_2 . The critical mixture curves for the $\text{N}(\text{C}_2\text{H}_5)_3/\text{CO}_2$ system were calculated by the Chueh–Prausnitz method¹⁵. The absence of a liquid phase was confirmed visually by use of a reactor equipped with sapphire windows. Similarly, in order to prevent the formation of a separate H_2 gas phase, the total pressure was kept at or above the mixture critical pressure for the H_2/CO_2 system¹⁶. The effect of the presence of $\text{N}(\text{C}_2\text{H}_5)_3$ on the phase behaviour of the H_2/CO_2 system is not known. After the required reaction time, the vessel was cooled to -196°C , vented, and then thawed. A sample of the product mixture, with an added internal standard, was dissolved in CD_3OD and analysed by ^1H NMR spectroscopy. The accuracy of this method was confirmed with known mixtures of HCOOH and $\text{N}(\text{C}_2\text{H}_5)_3$. The reaction was very clean and HCOOH was the sole product. The final yield after several hours was 3,700 moles HCOOH per mole of 1. The reaction is very rapid and attains a turnover frequency (TOF, the moles of product per mole catalyst per hour) of $1,400 \text{ h}^{-1}$ over the first hour. Figure 2 compares the turnover efficiencies of this straightforward reaction catalysed by several metal complexes in scCO_2 and ordinary solvents. The reaction catalysed by 1 at 50°C in scCO_2 is 18 times faster than in THF under otherwise identical conditions.

The catalysis with 1 proceeds smoothly without any induction

FIG. 1 Ruthenium-catalysed hydrogenation of carbon dioxide in the supercritical phase (scCO_2).

period. $\text{RuCl}_2(\text{P}(\text{CH}_3)_3)_4$ (compound 2 in Fig. 1) also efficiently catalyses the hydrogenation, although it shows a distinct induction period during which the yellow colour changes to colourless. The induction step is probably the amine-promoted slow conversion of 2 to 1 or $\text{RuHCl}(\text{P}(\text{CH}_3)_3)_4$ by analogy to the amine-promoted reaction of $\text{RuCl}_2(\text{P}(\text{C}_6\text{H}_5)_3)_3$ (refs 17–19). The actual chain-carrying catalysts could be some formate-Ru complexes^{20,21}. With 2, the production of formic acid attains equilibrium after 5 hours; with 1 we believe that true equilibrium

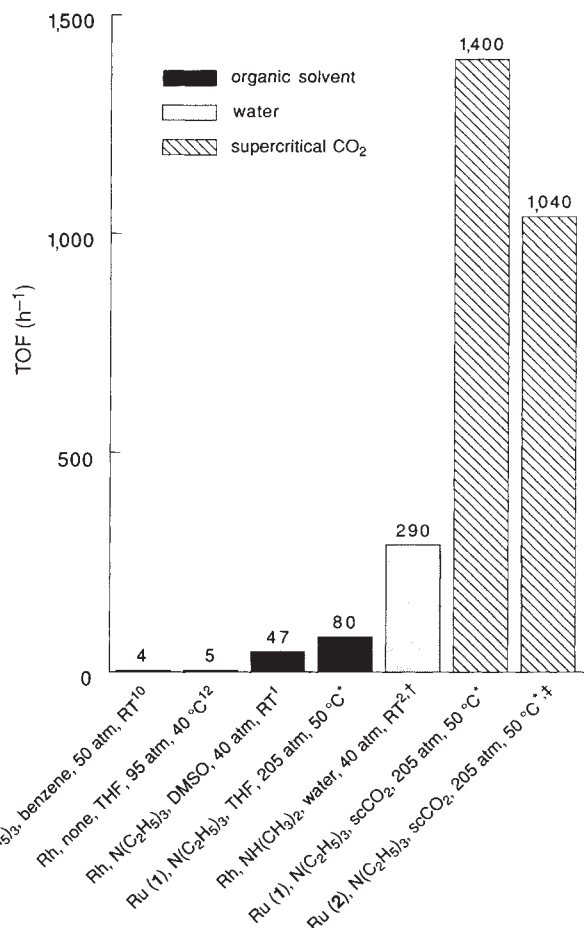


FIG. 2 The rates of formic acid production by CO_2 hydrogenation (TOF, turnover frequency, mol HCOOH per mol of catalyst per hour). Captions state the catalyst, base, solvent, total pressure and temperature (RT, room temperature). Rates were obtained by dividing published values of yield by reaction time (hours). (Source of data indicated by reference numbers; * this work. † Rate of reaction to equilibrium. Initial rate data unavailable. ‡ The rate refers to that after the induction period.)

is not attained. The greatest yields of formic acid have been obtained with catalyst **2**. The highest turnover number, 7,200 moles of formic acid per mole of catalyst, was recorded by using a larger reactor, a longer reaction time, and a catalyst : amine : water mole ratio of 1 : 12,000 : 30.

Excellent yields were obtained at H₂ pressures above 60 atm. The catalytic Ru species **2** is stable and maintains its high reactivity for a long time in scCO₂. The final yield of HCOOH corresponds to a HCOOH/N(C₂H₅)₃ mole ratio of up to 1.7, as was found in the DMSO system¹. This ratio decreased when a higher amine/Ru ratio was used; for the case of a catalyst : amine ratio of 1 : 12,000 mentioned above, the HCOOH : N(C₂H₅)₃ mole ratio is 0.6. This may be due to precipitation of liquid triethylammonium formate, which has poor solubility in scCO₂. Because the catalyst is soluble in both the scCO₂ and the newly forming liquid phase, the rate of reaction drops to a considerable extent. Triethylamine acts as a thermodynamic sink for the formic acid product; when it is saturated, the reaction stops. Thus, in the absence of the amine, no formic acid is produced. Use of a solid base such as potassium carbonate still allows production of HCOOH, although the TOF is reduced to only 9 h⁻¹.

RuH₂(P(C₆H₅)₃)₄, known to be the most active catalyst in benzene solution¹⁰, is inferior to **1** and **2** in scCO₂. We speculate

that the phenyl groups of this catalyst give it a lower solubility in scCO₂ than that enjoyed by **1**. Transition-metal complexes of trialkylphosphines are normally more soluble than those of triarylphosphines in nonpolar solvents; the higher solubility could be responsible for the higher activity. [Rh(norbornadiene)Cl]₂ with 1.4 eq. of bis(dicyclohexylphosphino)ethane as well as Pd on carbon were active but less so than **1** and **2**; the Rh complex may have decomposed.

The high performance of the formic acid production is not associated with a high concentration of CO₂. The reaction in liquid CO₂ containing **1**, water, and excess triethylamine under 205 atm (H₂ pressure of 85 atm) at 15 °C afforded a TOF of 1.3 h⁻¹; the TON and HCOOH/N(C₂H₅)₃ mole ratio were merely 20 and 0.013, respectively. Thus it is obvious that the present remarkable catalytic efficiency is based on the characteristics of scCO₂ including the extremely high miscibility with H₂ (refs 16, 22–24) and good mass-transfer capability.

This process is suited to a continuous flow system, with extraction and recycling of the amine and catalyst by scCO₂. Further technical refinements²⁵ are necessary, however, for the industrial realization of this high-pressure process.

Note added in proof: After this paper was submitted, another describing scCO₂ as both reactant and medium appeared²⁶. □

Received 4 November 1993; accepted 28 January 1994.

- Graf, E. & Leitner, W. *J. chem. Soc., chem. Commun.* 623–624 (1992).
- Gassner, F. & Leitner, W. *J. chem. Soc., chem. Commun.* 1465–1466 (1993).
- Saito, M., Yamauchi, Y. & Okuyama, T. (eds) *Fractionation by Packed Column SFC and SFE: Principles and Applications* (VCH, New York, in the press).
- Squires, T. G. & Paulaitis, M. E. (eds) *Supercritical Fluids: Chemical and Engineering Principles and Applications* (ACS Symp. Ser. No. 329, Am. chem. Soc., Washington DC, 1987).
- Shaw, R. W., Brill, T. B., Clifford, A. A., Eckert, C. A. & Franck, E. U. *Chem. Engng News* **69**, 26–39 (1991).
- Subramanian, B. & McHugh, M. A. *Industrial and Engineering Chemistry—Process Design and Development* **25**, 1–12 (1986).
- Aaltonen, O. & Rantakylä, M. *CHEMTECH* **21**, 240–248 (1991).
- Rathke, J. W., Klingner, R. J. & Krause, T. R. *Organometallics* **10**, 1350–1355 (1991).
- Weast, R. C. (ed.) *CRC Handbook of Chemistry and Physics* 70th edn (CRC, Boca Raton, Florida, 1990).
- Inoue, Y., Izumida, H., Sasaki, Y. & Hashimoto, H. *Chem. Lett.* 863–864 (1976).
- Darensbourg, D. J., Ovalles, C. & Pala, M. *J. Am. chem. Soc.* **105**, 5937–5939 (1983).
- Tsai, J.-C. & Nicholas, K. M. *J. Am. chem. Soc.* **114**, 5117–5124 (1992).
- Taqi Khan, M. M., Halligudi, S. B. & Shukla, S. J. *molec. Cat.* **57**, 47–60 (1989).

- Burgemeister, T., Kastner, F. & Leitner, W. *Angew. Chem. Int. Edn engl.* **32**, 739–741 (1993).
- Saito, M. & Nitta, T. in *Fractionation by Packed Column SFC and SFE: Principles and Applications* (eds Saito, M., Yamauchi, Y. & Okuyama, T.) (VCH, New York, in the press).
- Tsang, C. Y. & Streett, W. B. *Chem. Engng Sci.* **36**, 993–1000 (1981).
- Hallman, P. S., McGarvey, B. R. & Wilkinson, G. J. *chem. Soc., A* 3143–3150 (1968).
- Grushin, V. V. *Accs chem. Res.* **26**, 279–286 (1993).
- James, B. R., Markham, L. D., Hui, B. C. & Rempel, G. L. *J. chem. Soc., Dalton Trans.* 2247–2252 (1973).
- Komiyama, S. & Yamamoto, A. *J. organomet. Chem.* **46**, C58–C60 (1972).
- Kolominikov, I. S. et al. *J. organomet. Chem.* **59**, 349–351 (1973).
- Howdle, S. M. & Poliakov, M. J. *chem. Soc., chem. Commun.* 1099–1101 (1989).
- Howdle, S. M., Healy, M. A. & Poliakov, M. J. *Am. chem. Soc.* **112**, 4804–4813 (1990).
- Jobling, M., Howdle, S. M., Healy, M. A. & Poliakov, M. J. *chem. Soc., chem. Commun.* 1287–1290 (1990).
- Eggers, R. & Sievers, U. in *Supercritical Fluid: Science and Technology* (eds Johnston, K. P. & Penninger, J. M. L.) 478–498 (ACS Symp. Ser. No. 406, Am. chem. Soc., Washington DC, 1989).
- Reetz, M. T., Konen, W. & Strack, T. *Chimia* **47**, 493 (1993).

Platform subsidence behind an active subduction zone

T. A. Stern* & W. E. Holt†

* Research School of Earth Sciences, Victoria University of Wellington, Wellington, New Zealand

† Department of Earth and Space Sciences, State University of New York, Stony Brook, New York 11794-210, USA

NUMERICAL models of viscous mantle flow indicate that a broad subsidence should occur above subduction zones^{1–4}. This 'platform subsidence' is predicted to cause an asymmetrical depression 500–1,000 km wide, with a maximum depth of ~2 km if filled with water. Direct evidence linking platform subsidence to mantle flow has remained elusive, however. Here we offer evidence from western New Zealand for a complete link in both space and time between subduction initiation and platform subsidence. Because of the broad wavelength (>200 km) of the observed 1,500-m subsidence it is unlikely that the subsidence is due to a flexural foredeep. The observed subsidence does, however, match in both wavelength and amplitude that predicted by a model of subduction-induced mantle flow. The occurrence of platform subsidence has implications not only for understanding mantle convection, but also for explaining geoid anomalies over active margins⁵ and patterns of continental submergence⁶ and sedimentation⁷ that could otherwise be interpreted as a eustatic sea level change⁸.

Although it is recognized that platforms subside^{3,4,7} and that this subsidence is linked to subduction-induced mantle flow, there has been no example that provides a clear spatial and temporal correlation. In this study we establish such a correlation using subsidence data from a Miocene foreland basin. Taranaki Basin, on the western margin of central New Zealand, has wells distributed over ~170 km perpendicular to the Pacific–Australian plate boundary (Fig. 1). Up to 10 km of Cretaceous to Recent sediments are contained within Taranaki Basin⁹. The present plate boundary is east of the North Island (Fig. 1) and the northwest-dipping, subducted Pacific plate is now at depths of ~150–250 km beneath Taranaki Basin¹⁰. From Cretaceous to Oligocene times sedimentation was associated with rifting, opening and subsidence of the Tasman Sea¹¹. Around the Oligocene/Miocene boundary normal faults were inverted. Thrusting took place until mid-Miocene time in North Taranaki Basin and to the end of the Miocene period in South Taranaki Basin^{11,12}. This phase of compression has been linked to the onset of subduction at the eastern side of the North Island^{11,13}. Crustal thickening seen in deep seismic data, and Miocene thrusting, suggest that during the Early to mid-Miocene, southern Taranaki Basin developed as a flexurally controlled retro-arc foreland basin¹⁴ (see inset, Fig. 1).

Subsidence curves for 28 wells in Taranaki Basin (Fig. 1) were collated^{15,16} from palaeobathymetry and stratigraphic data. Tectonic subsidence (*Y*), which is the effective 'back-stripped' depth to basement in the absence of any sediment load, is a

function of the observed sedimentary thickness (corrected for compaction) (S) and the palaeobathymetry (W_d), and is given by:

$$Y = S[(\rho_m - \rho_s)/(\rho_m - \rho_w)] + W_d$$

where ρ_m , ρ_s , and ρ_w are the densities of mantle, sediments and water respectively¹⁷. Possible eustatic sea level changes are ignored as they would here be second order compared to changes in W_d .

For Taranaki Basin it is useful to analyse S and W_d separately because by the end of the Oligocene period much of New Zealand was submarine¹⁸ and sedimentation rates accordingly were low. At such times changes in tectonic subsidence (Y) are equal to changes in palaeobathymetry (W_d). For example, Fig. 2 shows the interaction of S and W_d for a well, Maui-2, in the middle of Taranaki Basin. A 1,500-m palaeobathymetric depression developed at ~25 Myr (before the present), but the principal sedimentation phase that filled the palaeobathymetric subsidence did not occur until ~17 Myr.

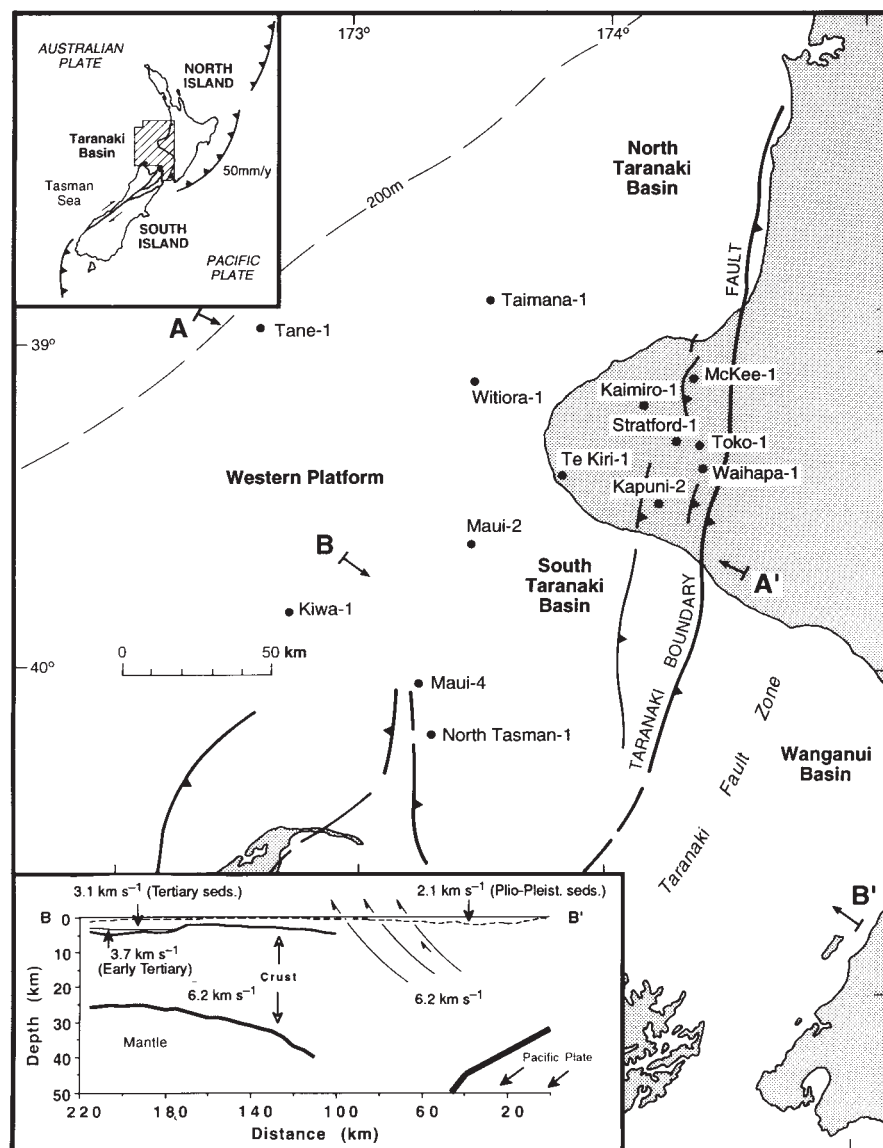
An initial interpretation¹⁴ of tectonic subsidence curves from Taranaki Basin was for a classical foreland basin model in terms

of a migrating flexural foredeep. Under this scheme, the palaeobathymetric depression (Fig. 2) at 25 Myr is the expression of isostatic loading by a developing fold and thrust belt in the adjacent Taranaki fault zone (Fig. 1), and the lack of sedimentation before 17 Myr is ascribed to much of New Zealand still being submarine. By 17 Myr the thrust belt emerged above sea level, leading to subaerial erosion and enhanced sedimentation (Fig. 2). This interpretation is, nevertheless, contradicted to some extent when we examine the relationship in both time and space for the development of the palaeobathymetric subsidence across Taranaki Basin.

Figure 3 shows the principal evidence for platform subsidence. A stack of palaeobathymetric profiles is shown for wells lying adjacent to a 170-km-long profile crossing the basin perpendicular to the Taranaki fault zone. A water-filled depression seems to have developed almost simultaneously across the basin at 25 ± 3 Myr. If the palaeobathymetry was driven by an advancing flexural foredeep, its position should be a function of time¹⁹.

An alternative method of presenting the data of Fig. 3 is given in Fig. 4a. Here we plot, west to east, the palaeobathymetry at two stages; at 37.5 and 19–17 Myr (upper and lower plots

FIG. 1 Location map of Taranaki Basin, Western Platform and oil-industry wells used in this study. A–A' refers to the cross-section shown in Fig. 4a, B–B' to the location of a deep seismic line whose depth-converted interpretation is summarized in the bottom inset¹⁴. Top inset, relationship between Australian and Pacific plates in the New Zealand region.



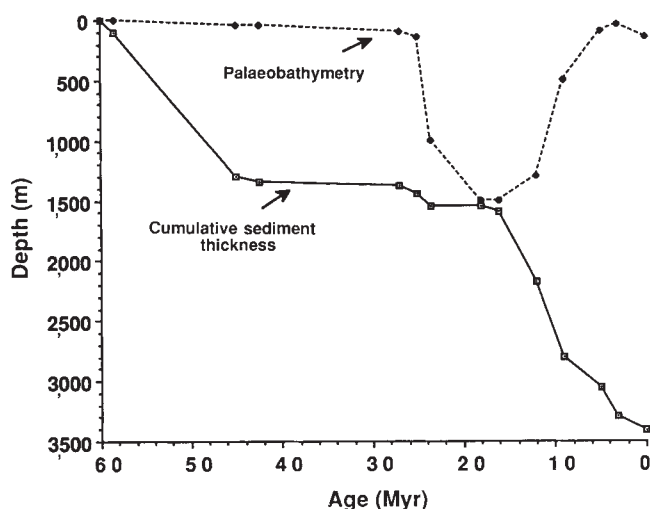


FIG. 2 Palaeobathymetry and cumulative sedimentary thickness data for the Maui-2 well¹⁵ (location shown in Fig. 1). Sedimentary thicknesses have been corrected for compaction effects¹⁵. Note 1,500 m of bathymetry at ~25 Myr yet the principal phase of sedimentation not occurring until 17 Myr.

respectively). This form of data presentation demonstrates the long wavelength of the subsidence. Figure 4b shows interpretation profiles plotted against the palaeobathymetric difference at 37.5 and 19 Myr from Fig. 4a. Two of the interpretation profiles are based on flexure due to loading of thrust sheets onto the edge of a discontinuous elastic lithospheric plate. Flexure profiles associated with two values of effective elastic thickness (T_e), 25 and 100 km, are shown because T_e is the principal parameter controlling the wavelength and amplitude of flexure²⁰. A value of $T_e = 25$ km is appropriate for Taranaki Basin as this has been previously determined for the Western Platform from a well-constrained loading study²¹. A T_e of 100 km represents an upper limit for continental lithosphere²⁰ and is well in excess of what we might anticipate for the recently rifted, and therefore thermally young, lithosphere of western New Zealand²¹.

Also shown in Fig. 4b are portions of two mantle-flow solutions for platform subsidence². These curves represent deflections of the Earth's surface above a subducted plate that sinks under its own excess weight through a highly viscous mantle². As the slab sinks, a viscous coupling of normal stresses from the slab

to the free surface of the Earth occurs, creating a low pressure region, and thus a broad platform subsidence^{2,4}. Theoretical curves² for water-filled deflections above slabs dipping at 30° and 60° are shown in Fig. 4b (mantle flow (1) and (2) respectively); these are similar in both amplitude and wavelength to other mantle-flow models^{3,4}. Overall, the mantle-flow models provide satisfactory fits to the long-wavelength palaeobathymetry whereas the elastic flexure models clearly do not.

A testable prediction of Gurnis's⁴ finite-element model of platform subsidence are a rapid (~500 m Myr⁻¹) subsidence immediately after subduction onset, and that this rate may continue for 10 Myr. Isostatic amplification of the tectonic depression by sediment loading is assumed in deriving this subsidence rate⁴. For water loading, which is the situation for the 11-Myr period between 28 and 17 Myr in Taranaki basin (Fig. 2), the predicted sediment-amplified subsidence rate should be divided

$$(\rho_m - \rho_w)/(\rho_m - \rho_s) = 2.4$$

for ρ_m , ρ_w , and ρ_s of 3,300, 1,000 and 2,400 kg m⁻³ (respectively). Thus a bathymetric subsidence rate of ~200 m Myr⁻¹ is

FIG. 3 A stack of palaeobathymetric curves for the wells that fall on or close to profile A-A' of Fig. 1. Note the simultaneous development of a bathymetric deep at 25 ± 3 Myr. Errors in assessing palaeobathymetry from foraminiferal faunas depend on the depth as shown by the error bars^{15,16}. Question marks indicate no data are available for Tane well between 37.5 and 25 Myr. Average subsidence rate over 28–17 Myr is 130 ± 30 m Myr⁻¹.

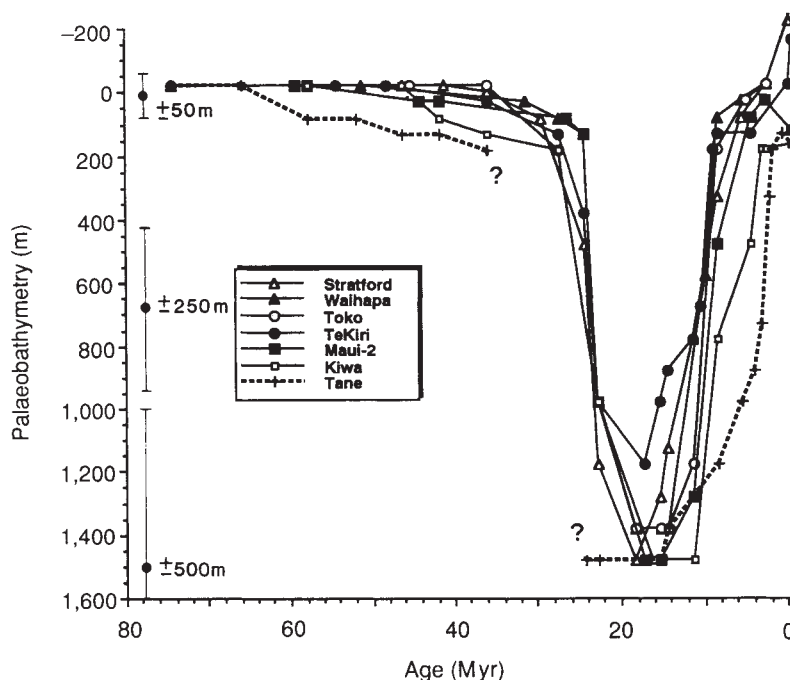
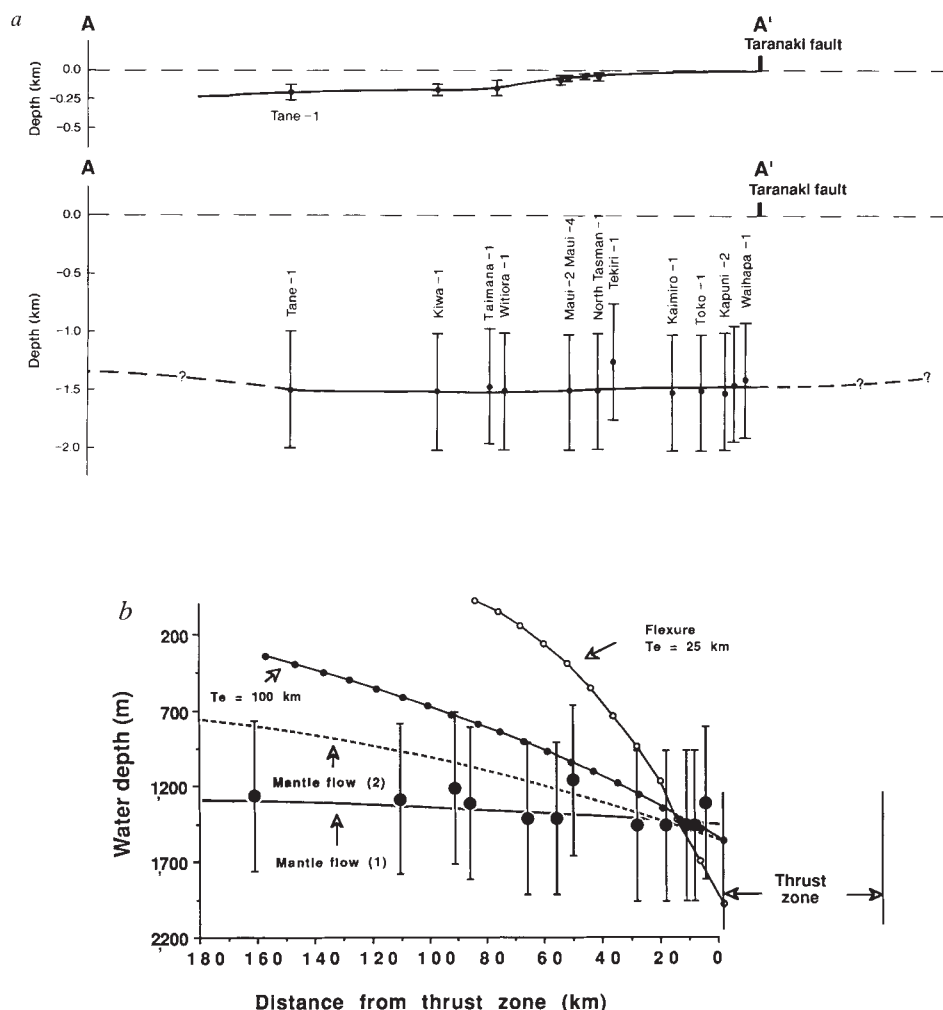


FIG. 4 a, Cross-sections of palaeobathymetry at 37.5 Myr (top) and 19 Myr (bottom) along profile A-A' in Fig. 1. b, Comparison of the difference between palaeobathymetry at 37.5 and 19–17 Myr with theoretical subsidence models driven by lithospheric flexure and mantle flow. Flexure profiles are for loading on the free edge of discontinuous elastic plates with effective elastic thicknesses (T_e) of 25 and 100 km. Loading was within a 50-km-wide zone at the free edge and constrained to produce a minimum 2-km-deep depression beneath the thrust belt. The flexure profiles were calculated with a finite-difference technique²⁷. Mantle-flow models are for 30° and 60° dipping slabs (Mantle flow 1 and 2 respectively)². The trench is assumed to be 200 km east of the Taranaki boundary fault. Water infilling of the induced subsidence and a mantle density of 3,300 kg m⁻³ are assumed for both the flexural and mantle-flow models.



predicted. An average subsidence rate of 130 ± 30 m Myr⁻¹ is observed for the period 28–17 Myr (Fig. 3). This is in reasonable agreement with the predicted rate.

Geological evidence of thrust faulting¹³ and onset of andesite volcanism²² from eastern and northern New Zealand points to subduction beneath the North Island of New Zealand starting at the beginning of the Miocene period. An Early Miocene age corresponds with our observation of a palaeobathymetric depression developing at 25 ± 3 Myr in Taranaki Basin. This link in timing, coupled with the character of the observed platform subsidence (Fig. 4b), are the two principal observations leading us to ascribe the long-wavelength Taranaki Basin subsidence to

mantle flow driven by the sinking of a subducted plate.

We suggest that platform subsidence and flexure, caused by thrust-sheet loading, have operated simultaneously within Taranaki Basin²³. Such a dual process has been suggested elsewhere³ and, if general, would explain some previously documented problems^{24–26} associated with interpreting gravity and isopach data from foreland basins. In particular, the difficulty in identifying the complete driving load for creating and maintaining foreland basins, sometimes referred to as the 'hidden load problem' may be explicable in terms of the temporal and spatial overlapping of both platform subsidence and isostatic subsidence due to thrust-sheet loading. □

Received 18 August 1993; accepted 14 February 1994.

- McKenzie, D. P. *Geophys. J. R. astr. Soc.* **18**, 1–32 (1969).
- Harper, J. F. *Phys. Earth planet. Inter.* **36**, 285–290 (1984).
- Mitrovica, J. X., Beaumont, C. & Jarvis, G. T. *Tectonics* **8**, 1079–1094 (1989).
- Gurnis, M. *Science* **255**, 1556–1558 (1992).
- Davies, G. F. *Earth planet. Sci. Lett.* **54**, 431–441 (1981).
- Gurnis, M. *Geology* **21**, 29–32 (1993).
- Cross, T. A. & Pilger, R. H. *Nature* **274**, 653–657 (1978).
- Bond, G. *Geology* **4**, 557–560 (1976).
- Thrasher, G. P. & Cahill, J. P. New Zealand Geological Survey Report G-142 (Dept of Scient. & Ind. Res., Wellington, New Zealand, 1989).
- Adams, R. D. & Ware, D. E. N. Z. *J. Geol. Geophys.* **20**, 59–83 (1976).
- King, P. R. & Thrasher, G. P. *Mem. Am. Ass. Petrol. Geol.* **53**, 93–118 (1992).
- Pilaar, W. F. H. & Wakefield, L. L. *J. Aust. Petrol. Explor. Assoc.* **18**, 93–101 (1978).
- Rait, G., Chanier, F. & Waters, D. W. *Geology* **19**, 230–233 (1991).
- Stern, T. A. & Davey, F. J. *Geology* **18**, 979–982 (1990).

- Hayward, B. W. & Wood, R. A. New Zealand Geological Survey Report, PAL 147 (Dept of Scient. & Ind. Res., Lower Hutt, 1989).
- Hayward, B. W. *J. foram. Res.* **20**, 71–83 (1990).
- Steckler, M. S. & Watts, A. B. *Earth planet. Sci. Lett.* **41**, 1–13 (1978).
- Carter, R. M. & Norris, R. J. *Earth planet. Sci. Lett.* **31**, 85–94 (1976).
- Beaumont, C. *Geophys. J. R. astr. Soc.* **65**, 291–329 (1981).
- Watts, A. B. *Basin Res.* **4**, 169–178 (1992).
- Holt, W. E. & Stern, T. A. *Earth planet. Sci. Lett.* **107**, 523–538 (1991).
- Ballance, P. F. *Earth planet. Sci. Lett.* **28**, 356–370 (1976).
- Holt, W. E. & Stern, T. A. *Tectonics* (in the press).
- Karner, G. D. & Watts, A. B. *J. geophys. Res.* **88**, 10449–10477 (1983).
- Nunn, J. A., Czernik, M. & Pilger, R. H. *Tectonics* **6**, 603–617 (1987).
- Flemings, P. B. & Jordon, T. E. *J. geophys. Res.* **94**, 3851–3866 (1989).
- Bodine, J. H., Steckler, M. S. & Watts, A. B. *J. geophys. Res.* **86**, 3695–3707 (1981).

ACKNOWLEDGEMENTS. We thank J. Harper, R. Walcott and P. Barrett of Victoria University for reviews of an initial draft of the manuscript.

The endogenous oestrogen metabolite 2-methoxyoestradiol inhibits angiogenesis and suppresses tumour growth

Theodore Fotsis*, Youming Zhang†, Michael S. Pepper‡, Herman Adlercreutz§, Roberto Montesano‡, Peter Paul Nawroth† & Lothar Schweigerer*

* Department of Oncology and Haematology, Children's University Hospital, University of Heidelberg, INF 150, 69120 Heidelberg, Germany

† Department of Medicine and Pathology, University of Heidelberg, 69120 Heidelberg, Germany

‡ Institute of Histology and Embryology, Department of Morphology, University Medical Center, 1121 Geneva 4, Switzerland

§ Department of Clinical Chemistry, University of Helsinki, Meilahti Hospital, SF-00290 Helsinki, Finland

THE formation of new blood vessels (angiogenesis) is critical for the growth of tumours¹⁻³ and is a dominant feature in various angiogenic diseases such as diabetic retinopathy, arthritis, haemangiomas and psoriasis⁴. Recognition of the potential therapeutic benefits of controlling pathological angiogenesis has led to a search for angiogenesis inhibitors. Here we report that 2-methoxyoestradiol, an endogenous oestrogen metabolite of previously unknown function, is a potent inhibitor of endothelial cell proliferation and migration as well as angiogenesis *in vitro*. Moreover, when administered orally in mice, it strongly inhibits the neovascularization of solid tumours and suppresses their growth. Unlike the angiostatic steroids of corticoid structure⁵, it does not require the co-administration of heparin or sulphated cyclodextrins for activity. Thus, 2-methoxyoestradiol is the first steroid to have high antiangiogenic activity by itself. Our results suggest that this compound may have therapeutic potential in cancer and other angiogenic diseases.

We have shown previously that fractionated human urine contains substances capable of inhibiting cell proliferation and angiogenesis^{6,7}. Two of the most potent fractions contained catecholic and non-polar oestrogen metabolites, suggesting their involvement in the oestrogen-dependent regulation of growth

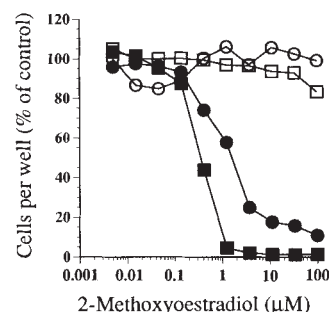


FIG. 1 Effect of increasing concentrations of 2-methoxyoestradiol on the proliferation of endothelial cells or fibroblasts seeded at various densities. Bovine brain capillary endothelial cells (squares) or human skin fibroblasts (HFK2) (circles) were suspended in DMEM containing 10% newborn or fetal calf serum, respectively, and seeded in 1-ml aliquots containing either 1.5×10^4 (■, ●) or 1×10^5 (□, ○) cells into 12-well dishes. BBCE cells seeded at 1.5×10^4 per ml received bFGF (2.5 ng ml^{-1}) every other day plus either buffer only (controls) or the indicated concentrations of 2-methoxyoestradiol; HFK2 cells seeded at 1.5×10^4 per ml received either buffer only (controls) or 2-methoxyoestradiol. When seeded at a density of 1×10^5 per ml, BBCE or HFK2 cells were grown to confluence. The medium was then changed and the cells received either buffer (controls) or the indicated concentrations of 2-methoxyoestradiol. In all experiments, substances were added in 10-μl aliquots and cells were counted after 6 days. Values are expressed as per cent of control (cells receiving no 2-methoxyoestradiol) and are the means of duplicate determinations which varied by <10% of the mean.

and angiogenesis. We therefore examined synthetic standards of these and other related oestrogen metabolites for their effects on basic fibroblast growth factor (bFGF)-induced proliferation of bovine brain capillary endothelial cells (BBCE) *in vitro* (Table 1). 2-Methoxyoestradiol showed the strongest inhibitory effect of all metabolites tested, the half-maximal inhibitory concentration (IC_{50}) being $\sim 0.15 \mu\text{M}$ (Table 1 and Fig. 1). The inhibition seemed to be specific to the 2-methoxyoestradiol molecule, as isomeric and closely related structures were 40–250 times less potent, with some oestrogen metabolites having no effect at all.

The inhibitory effect of 2-methoxyoestradiol was essentially the same (data not shown) on different types of endothelial cells (bovine aortic, bovine adrenal cortex microvascular (BME),

TABLE 1 Effect of 2-methoxyoestradiol and related structures on the proliferation of endothelial cells

Trivial name	Systematic name	IC_{50} (μM)
2-Methoxyoestradiol	1,3,5 (10)-Oestratriene-2,3,17β-triol 2-methyl ether	0.134
2-Hydroxyoestrone	1,3,5 (10)-Oestratriene-2,3-diol-17-one	5.9
4-Methoxyoestradiol	1,3,5 (10)-Oestratriene-3,4,17β-triol 4-methyl ether	7.24
2-Methoxyoestradiol 3-methyl ether	1,3,5 (10)-Oestratriene-2,3,17β-triol 2,3-dimethyl ether	7.78
2-Methoxyoestrone	1,3,5 (10)-Oestratriene-2,3-diol-17-one 2-methyl ether	11.5
2-Hydroxyoestradiol	1,3,5 (10)-Oestratriene-2,3,17β-triol	15.7
2-Methoxyoestriol	1,3,5 (10)-Oestratriene-2,3,16α, 17β-tetrol 2-methyl ether	21.3
2-Hydroxyoestradiol 3-methyl ether	1,3,5 (10)-Oestratriene-2,3,17β-triol 3-methyl ether	21.42
Oestrone	1,3,5 (10)-Oestratriene-3-ol-17-one	26
16-Epioestriol	1,3,5 (10)-Oestratriene-3,16β, 17β-triol	31
16α-Hydroxyoestrone	1,3,5 (10)-Oestratriene-2,16α-diol-17-one	31.1
Oestriol	1,3,5 (10)-Oestratriene-3,16α,17β-triol	32
Oestradiol-17β	1,3,5 (10)-Oestratriene-3,17β-diol	34.5
15α-Hydroxyoestriol	1,3,5 (10)-Oestratriene-3,15α,16α,17β-tetrol	NI
Oestriol-16α-glucuronide	1,3,5 (10)-Oestratriene-3,16α,17β-triol 16α-glucosiduronate	NI

Bovine brain-derived capillary endothelial cells were adjusted to a density of 5,000 cells per ml in DMEM plus 10% newborn calf serum and seeded in 1-ml aliquots per well into 12-well cluster dishes. The cells received 10-μl aliquots of bFGF (2.5 ng ml^{-1} medium) every other day and either buffer (10 μl) or buffer containing increasing concentrations of 2-methoxyoestradiol or the other substances to be tested. Cells were counted after 6 days with a Coulter particle counter. The results are expressed as IC_{50} values, that is, the concentration of the respective substance resulting in half the numbers of the cells that were obtained in controls (cells receiving bFGF and buffer only, usually in the range of 200,000 cells per well). Values are the means of duplicate determinations which varied by <10% of the mean. NI, no inhibition.

human umbilical vein and human skin), with IC_{50} values varying between 0.2 and 2.5 μM . 2-Methoxyoestradiol reversibly inhibited endothelial cell proliferation at concentrations up to 3–5 μM , being increasingly cytotoxic (reduction of cell number below the initial seeding density) at higher doses (data not shown). Confluent cultures of endothelial cells were, however, virtually unaffected by 2-methoxyoestradiol; marginal cytotoxicity was observed at 100 μM (Fig. 1). Interestingly, when treated

simultaneously with bFGF, confluent cultures of endothelial cells unable to further divide exhibited cytotoxicity similar to low-density cultures (data not shown). Therefore, 2-methoxyoestradiol seems to target only cells stimulated by bFGF or other growth factors. In addition to the effects on proliferation, 2-methoxyoestradiol had effects on other functions of endothelial cells important for angiogenesis. It inhibited the *in vitro* migration of BBCE and BME cells (data not shown) and the ability

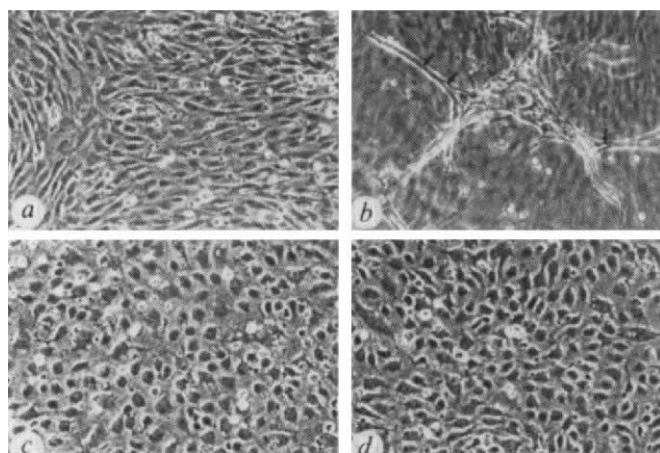
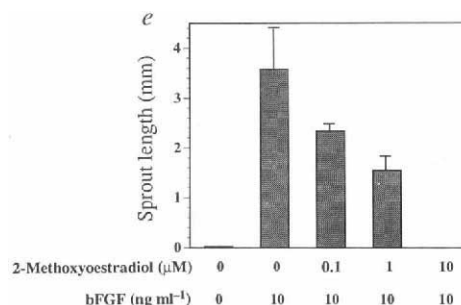


FIG. 2 Effect of 2-methoxyoestradiol on *in vitro* angiogenesis. Bovine microvascular endothelial cells were grown to confluence on three-dimensional collagen gels⁸. After confluence, they received either no additions (a), bFGF (10 ng ml⁻¹) only (b), 2-methoxyoestradiol (10 μM) only (c), or 2-methoxyoestradiol (10 μM) 2 h before bFGF (10 ng ml⁻¹) (d). Fresh medium and compounds were added every 2–3 days, and the cultures were photographed and fixed after 7 days (magnification $\times 125$). In b, the lumina of the tubules are evident as



translucent slits under phase contrast microscopy (arrows); in d, 2-methoxyoestradiol inhibits the formation of such tubes. In c, 2-methoxyoestradiol alone causes a more polygonal morphology of endothelial cells than the culture in a. e, The results shown in a–d were analysed quantitatively by photographing three randomly selected fields per well measuring 1×1.4 mm, each at a single level beneath the surface monolayer¹⁴. Invasion was quantified by measuring the total additive length of all cell cords that had penetrated the underlying gel. Results are expressed as the mean length in mm $\pm$ s.e.m. per field from at least three photographic fields per well. A single representative experiment is shown; this experiment has been repeated four times. 2-Methoxyoestradiol at 0.1, 1 and 10 μM reduced the bFGF-induced invasion of BME by 35, 57 and 100%, respectively, with an IC_{50} of ~ 0.35 μM .

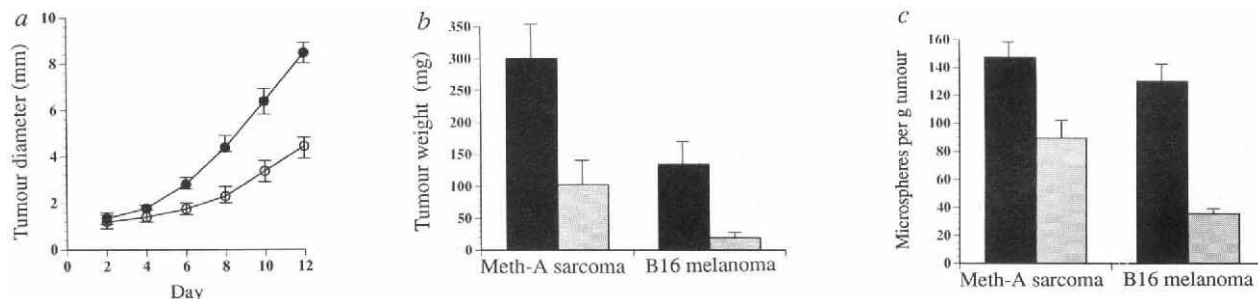
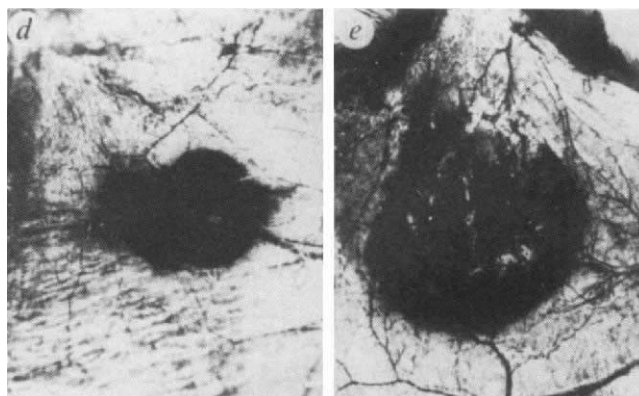


FIG. 3 Inhibition of growth (a, b) and neovascularization (c–e) of tumours by oral administration of 2-methoxyoestradiol in mice. a, b, Meth-A sarcoma or B16 melanoma cells (1×10^6) were inoculated subcutaneously in 0.1 ml saline in the dorsal skin of 20 C3H mice. On the same day, half the mice received orally 2-methoxyoestradiol at 100 mg per kg body weight, suspended in 300 μl olive oil, and half received 300 μl olive oil alone. This treatment was carried out every day and the diameter of the tumours was monitored every second day. On day 12 the mice were killed and the tumours weighed. a, Shows the time course of the growth of the Meth-A sarcoma tumours, expressed as mean diameter in mm $\pm$ range (○, 2-methoxyoestradiol-treated; ●, control), while b, shows the weight in mg $\pm$ s.d. of both the Meth-A sarcoma and B16 melanoma tumours on day 12 (□, 2-methoxyoestradiol-treated; ■, control). c, In six control (■) and six 2-methoxyoestradiol-treated (□) mice the neovascularization of the Meth-A sarcoma and B16 melanoma tumours was quantified by injecting 5×10^5 microspheres (10 μm , E-Z Trac, Interactive Medical Technology, Los Angeles, CA) into the left ventricle 5 min before the mice were killed. After the tumours had been weighed (see above), the tumour tissue was dissolved with a sodium hydroxide–SDS solution and the microspheres were counted microscopically. In c, the degree of neovascularization of the treated and control tumours is expressed as microspheres per g tumour tissue $\pm$ s.d. d, e, In a few control and 2-methoxyoestradiol-treated mice, the dorsal



skin, together with the B16 melanoma tumour, was excised and the blood vessels in the subcutaneous fascia in the treated (d) and control (e), mice were visualized with Indian ink. d, $\times 15$; e, $\times 20$.

of BME cells to invade collagen gels and form capillary-like structures⁸, at concentrations comparable to those that inhibited *in vitro* proliferation of endothelial cells (Fig. 2a–e); the IC₅₀ of *in vitro* angiogenesis was ~0.35 µM (Fig. 2e).

The inhibitory effects of 2-methoxyoestradiol were not restricted to endothelial cells. The proliferation of low-density cultures of various normal (bovine granulosa, NIH3T3 mouse embryonic fibroblast and HFK2 human skin fibroblast) and tumour (SH-EP human neuroblastoma, A204 human rhabdomyosarcoma and Y-79 human retinoblastoma) cells were inhibited, with IC₅₀ concentrations of 0.35–2.2 µM (Fig. 1 and data not shown). Again, quiescent confluent monolayers of HFK2 cells were virtually unaffected at doses up to 100 µM (Fig. 1). These results indicate that actively proliferating cells are the target of 2-methoxyoestradiol, and that the compound probably interferes with a proliferation-associated cellular event.

The *in vitro* effects of 2-methoxyoestradiol on angiogenesis and proliferation of malignant cells prompted us to investigate its *in vivo* properties. When administered orally in mice, 2-methoxyoestradiol inhibited the growth of tumours arising from subcutaneously injected Meth A sarcoma and B16 melanoma cells (Fig. 3a, b). This tumour-suppressing effect seemed to be due to inhibition of tumour-induced angiogenesis rather than direct inhibition of the proliferation of tumour cells. This was assessed by quantifying the vasculature of tumours using microspheres injected in the left ventricle shortly before the mice were killed and the tumours excised. After the tumours had been weighed, the tumour tissue was dissolved and the microparticles were counted. There was a significant reduction in the number of microspheres per g tumour tissue in the treated compared with the control tumours (Fig. 3c); the result was confirmed by visualization of the capillaries with Indian ink (Fig. 3d, e). Apart from their marginally lower weight (~15%), the treated mice showed no apparent signs of toxicity and were all alive after 12 days of daily treatment. We have not observed hair loss, intestinal disturbance or infection, all toxic side-effects associated with conventional chemotherapy. The lack of toxicity is consistent with the *in vitro* results showing no effect on quiescent, non-dividing cells (Fig. 1). The marked difference in the tumour weight of treated and control animals (75.7%) cannot be explained by the 15% weight loss of the mice.

Unlike the angiostatic steroids of corticoid structure⁵, heparin or sulphated cyclodextrins are not required for the antiangiogenic activity of 2-methoxyoestradiol, indicating a different mechanism of action. The results in Table 1 and the negligible affinity of 2-methoxyoestradiol for the oestrogen receptor⁹ also exclude the possibility that the effects of this steroid on endothelial cells are mediated through interactions with the oestrogen receptor. We have observed, by immunofluorescence, disruption of microtubules, but not actin microfilaments, vimentin intermediate filaments or vinculin-containing adhesion plaques in endothelial cells treated with 2-methoxyoestradiol (data not shown); this suggests that abnormal microtubule assembly might be responsible for its effects on proliferating cells^{10,15}. We have also observed that 2-methoxyoestradiol reduces the bFGF-induced increase in urokinase-type plasminogen activator activity in BME cells, without affecting levels of activity of tissue-type plasminogen activator or plasminogen activator inhibitor-1; this effect is partially mimicked by colchicine and other microtubule-disrupting agents (M.S.P. *et al.* manuscript in preparation). Because increased endothelial cell proteolysis is required for extracellular matrix invasion during angiogenesis, the selective reduction in urokinase-type plasminogen activator activity may be responsible in part for the antiangiogenic effect of 2-methoxyoestradiol.

Antiangiogenic therapy of human diseases is a promising new concept that has already been applied in the treatment of refractory haemangiomas with α -interferon^{11,12} and nerve-sheath tumours with AGM-1470 (ref. 13). The *in vitro* and *in vivo* inhibitory effects of 2-methoxyoestradiol on angiogenesis suggest

that this compound is a novel antiangiogenic therapeutic agent which could be used in the treatment of solid tumours and other angiogenic diseases such as paediatric haemangiomas, rheumatoid arthritis, psoriasis and diabetic retinopathy. □

Received 17 November 1993; accepted 7 January 1994.

1. Folkman, J., Watson, K., Ingber, D. E. & Hanahan, D. *Nature* **339**, 58–61 (1989).
2. Blood, C. H. & Zetter, B. R. *Biochim. biophys. Acta Rev. Cancer* **1032**, 89–118 (1990).
3. Mahadevan, V. & Hart, I. R. *Eur. J. Cancer* **27**, 679–680 (1991).
4. Klagsbrun, M. & Folkman, J. in *Peptide Growth Factors and Their Receptors II* (eds Sporn, M. B. & Roberts, A. B.) 549–586 (Springer, Berlin, 1990).
5. Crum, R., Szabo, S. & Folkman, J. *Science* **230**, 1375–1378 (1985).
6. Fotis, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **90**, 2690–2694 (1993).
7. Schweigerer, L. *et al.* *Eur. J. clin. Invest.* **22**, 260–264 (1992).
8. Montesano, R., Vassalli, J.-D., Baird, A., Guillemin, R. & Orci, L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7297–7301 (1986).
9. MacLusky, N. J., Barnea, E. R., Clark, C. R. & Naftolin, F. *Catechol Estrogens* (eds Merriam, G. R. & Lipsett, M. B.) 151–165 (Raven, New York, 1983).
10. Seegers, J. C. *et al.* *J. Steroid Biochem.* **32**, 797–809 (1989).
11. White, C. W., Sondheimer, H. M., Crouch, E. C., Wilson, H. & Fan, L. L. *New Engl. J. Med.* **320**, 1197–1200 (1989).
12. Ezekowitz, A., Mulliken, J. & Folkman, J. *Br. J. Haemat.* **79** (suppl. 1), 67–68 (1991).
13. Takamiya, Y., Friedlander, R. M., Brem, H., Malick, A. & Martuza, R. L. *J. Neurosurg.* **78**, 470–476 (1993).
14. Pepper, M. S., Ferrara, N., Orci, L. & Montesano, R. *Biochem. biophys. Res. Commun.* **189**, 824–831 (1992).
15. D'Amato, R. J., Lin, C. M., Flynn, E., Folkman, J. & Hamel, E. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

ACKNOWLEDGEMENTS. We thank G. Schneckenburger, C. Di Sanza and M. Quayzin for technical assistance, B. Favri for photography, M. B. Furie and S. C. Silverstein for BME cells, W. Franke for cytoskeletal staining, and A. Habenicht for his generous encouragement to T.F. and L.S. This work was supported by grants from Schwerpunkt 'Entzündung' of the Land Baden-Württemberg, Verein zur Förderung der Krebsforschung in Deutschland (e.V.), Deutsche Leukämieforschungshilfe-Aktion für Krebskranke Kinder (e.V.) Heidelberg, Deutsche Forschungsgemeinschaft and Deutsche Krebshilfe (T.F., Y.Z., P.P.N. and L.S.), the Swiss NSF (M.S.P. and R.M.), and the MRC of the Academy of Finland (H.A.).

Facial shape and judgements of female attractiveness

D. I. Perrett*, K. A. May* & S. Yoshikawa†

* School of Psychology, University of St Andrews, Fife, KY16 9JU, UK

† Department of Psychology, Faculty of Letters, Ottemon Gakuin University, Ibaraki, Osaka 567, Japan

THE finding that photographic^{1–4} and digital⁵ composites (blends) of faces are considered to be attractive has led to the claim that attractiveness is averageness⁵. This would encourage stabilizing selection, favouring phenotypes with an average facial structure⁵. The 'averageness hypothesis' would account for the low distinctiveness of attractive faces⁶ but is difficult to reconcile with the finding that some facial measurements correlate with attractiveness^{7,8}. An average face shape is attractive but may not be optimally attractive⁹. Human preferences may exert directional selection pressures, as with the phenomena of optimal outbreeding and sexual selection for extreme characteristics^{10–14}. Using composite faces, we show here that, contrary to the averageness hypothesis, the mean shape of a set of attractive faces is preferred to the mean shape of the sample from which the faces were selected. In addition, attractive composites can be made more attractive by exaggerating the shape differences from the sample mean. Japanese and Caucasian observers showed the same direction of preferences for the same facial composites, suggesting that aesthetic judgements of face shape are similar across different cultural backgrounds. Our finding that highly attractive facial configurations are not average shows that preferences could exert a directional selection pressure on the evolution of human face shape.

The averageness hypothesis was tested directly by comparing the attractiveness of composite faces that had different shapes but identical skin textures. A composite image made from many individuals of a group should maintain any consistent visual attributes possessed by that group. If attractiveness is an index

of averageness, a composite with a shape derived from highly attractive faces should not differ from a composite with the average face shape. Alternatively, if attractive faces are systematically different from average, a composite with a shape derived from highly attractive faces should be preferred. A further prediction of the averageness hypothesis is that, if any face shape is exaggerated away from average, it should become less attractive.

These predictions were first examined with a collection of 60 caucasian female face images. The images, presented in random order, were rated for attractiveness on a 7-point Likert scale (1 very unattractive, 7 very attractive) by caucasian subjects (26 female, age 18–45 yr, and 10 male, age 19–24 yr). Ratings were

correlated across subject gender (Spearman's $r_{(58)} = 0.75$, $P < 0.00005$) and were consistent across subjects (mean inter-rater correlation $r_{(58)} = 0.41$, $P < 0.005$; estimated by correlating each subject's ratings with those of one other randomly chosen subject of the same gender and averaging standard deviate (z) score transformed correlation coefficients).

The 'average' face shape was defined as the mean shape of the entire sample (see Fig. 1 for methods). The 'high' face shape was derived from the 15 faces with highest attractiveness ratings. The high shape was then caricatured by exaggerating the shape differences from the average by 50% to create a 'high + 50%' shape. Composite images with average, high and high + 50%



FIG. 1 *a–c*, Shape of average, attractive and caricature of attractive caucasian female faces. Images of 60 UK caucasian female faces (age 20–30 yr; posing with neutral expression) were frame-grabbed in 24-bit colour (531 horizontal by 704 vertical pixel resolution). The shape of facial features of the 60 faces was defined by manually marking 224 predefined feature points (for example, tip of the nose) for each digital face image¹⁸. The average position of each feature point was calculated for the entire sample of 60 faces and the appropriate points were joined to produce a line-drawn representation of the 'average' face shape (*a*). The mean shape of attractive faces ('high' shape; *b*) was defined by averaging the positions of the feature points for the 15 faces with highest attractiveness ratings. To caricature¹⁸ any distinctive attributes of the high shape, the differences between the 224 feature positions in the high and average shapes were increased by 50% to define the 'high + 50%' shape (*c*). The differences in feature positions between

average and high + 50% shapes are indicated in *c* as superimposed vector differences calculated after scaling, rotating and translating these two shapes so that the pupils were aligned. *d–f*, Composite images. Each of the 60 original images was distorted¹⁹ (that is, 'warped' or 'morphed') into the 'average' shape. To achieve this, the feature positions defining the shape of each face were used to divide original images into triangular tessellations. Images were distorted by taking each individual triangle and remapping the enclosed pixels to fit the shape of the corresponding triangle in the tessellation structure for the average shape. The 60 images with the same remapped shape were then blended together²⁰ to form the average composite (*d*) by averaging the colours and intensities of corresponding pixels in constituent images. High and high + 50% composite images (*e* and *f*) were similarly obtained by remapping the 60 original images into the high and high + 50% shapes before blending.

shapes, but identical skin textures, were generated from the 60 original faces (see Fig. 1d–f for methods).

A new set of 36 UK caucasian subjects (age 22–46 yr; 18 female, 18 male) was given pairs of composites and asked to choose the most attractive in each pair. Male subjects preferred the high to the average composite (16/18 subjects; binomial test, $P < 0.001$) and preferred the high + 50% to the high composite (15/18; $P < 0.005$). Female subjects expressed the same preferences (15/18, $P < 0.005$; 14/18, $P < 0.02$, respectively). These results show that attractiveness is not averageness: first, the high composite was preferred over the average; second, when the high composite was caricatured to increase the difference from average, the attractiveness increased.

To examine the generality of the effects across faces, we extended the study to a set of Japanese female faces rated for attractiveness by Japanese subjects. Average, high and high + 50% composites were created using similar methods (Fig. 2). A new group of 26 Japanese subjects (undergraduates in Japan, aged 19–29 yr; 9 female, 17 male) preferred the high to the average composite (21/26 subjects, $P < 0.01$) and preferred the high + 50% to the high composite (19/26, $P < 0.02$). Thus, exaggeration of the differences between the high and average shapes again enhanced attractiveness.

To examine the cross-cultural consistency of the manipulations, the composite Japanese faces were shown to 36 caucasian subjects (undergraduates in the UK, aged 19–26 yr; 25 female, 11 male). These subjects also preferred the high to the average composite (25/36, $P < 0.02$) and preferred the high + 50% to the high composite (25/36, $P < 0.02$). Thus, caucasian and Japanese subjects showed the same pattern of preferences with the same face stimuli. This is consistent with previous findings of greater similarities than differences in cross-cultural judgements of facial attractiveness^{15,16}.

Figure 1c shows the differences in physical dimensions between average and attractive (high + 50%) caucasian female face shapes. The more attractive shape had higher cheek bones, a thinner jaw and larger eyes relative to the size of the face than the average shape. The attractive shape also had shorter distances between the mouth and chin and between nose and mouth. Average and attractive (high + 50%) Japanese face shapes showed similar differences in the eyes, mouth and chin (Fig. 2c). The Japanese face shapes also showed more pronounced differences in eyebrow shape/position and an additional change in height on the nasal side of the cheek.

We found similar preferences for a non-average shape with male faces. High and average composites were made with methods described in Fig. 1 from a set of 59 UK caucasian male faces (age 20–30 yr) initially rated for attractiveness by UK caucasian subjects (age 18–30 yr; 15 male, 15 female; ratings correlated across gender of rater, $r_{(57)} = 0.92$, $P < 0.0001$, and across individual rater, estimated mean correlation $r_{(57)} = 0.45$, $P < 0.005$). A new set of caucasian subjects (age 18–30 yr; 18 male, 18 female) preferred the high to the average composite (14/18 male subjects, $P < 0.02$; 13/18 female subjects, $P < 0.05$). The attractiveness of the high composite was not enhanced by caricaturing.

Thus, our results show that highly attractive faces are systematically different in shape from average. The similarity of attractive facial characteristics across two cultures is consistent with the claim that such characteristics are functionally significant. Viewpoints vary on the functional significance of attractiveness in humans^{7,8,17} and preferences in general^{5,10–14}. Attractive facial features may signal sexual maturity and fertility, emotional expressiveness or a 'cuteness' generalized from parental protectiveness towards young^{7,8}. If human facial attractiveness affects fitness in an evolutionary sense, then the preference for non-

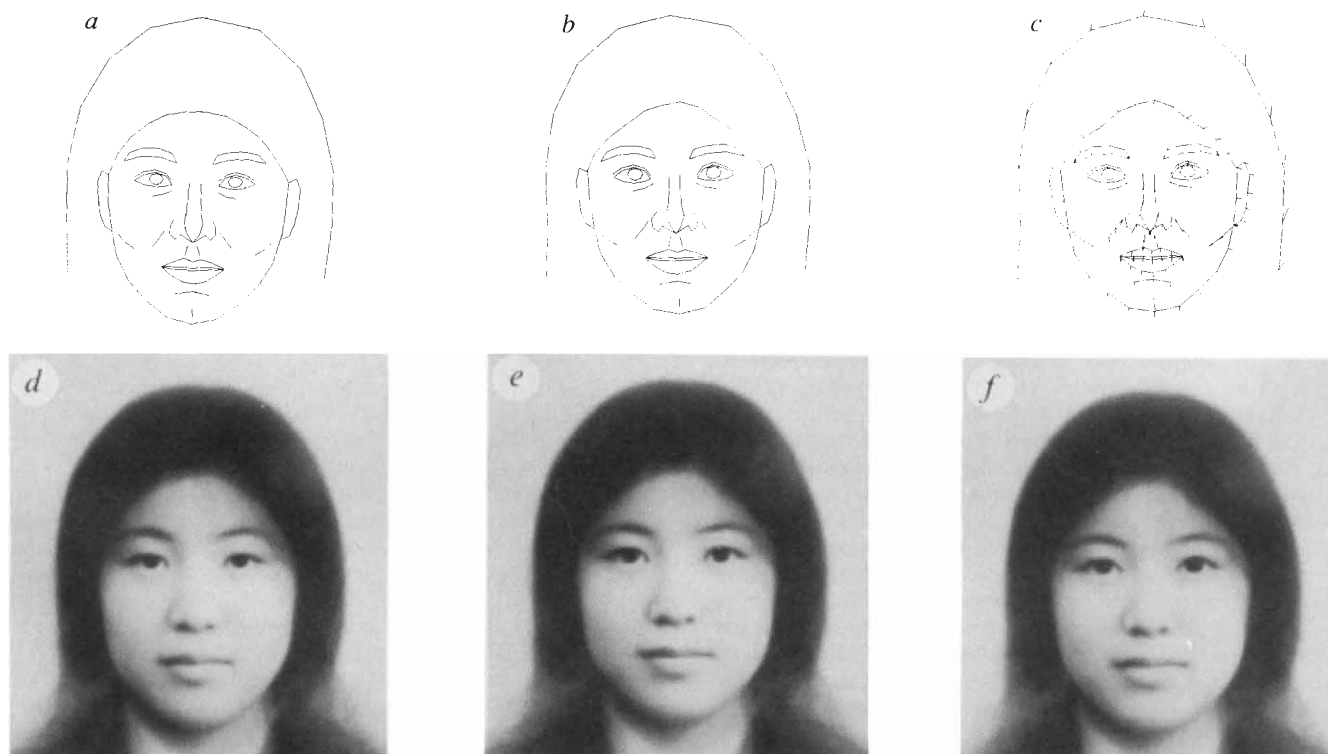


FIG. 2 Shapes and composite images of Japanese faces. The faces of 342 Japanese high-school girls (age 18–19 yr; posing with neutral expression) were photographed and rated for attractiveness on a 7-point Likert scale by 12 Japanese subjects (age 23–28 yr; 6 female, 6 male). The photographs were frame-grabbed in 8-bit grey-scale format. Using methods described in Fig. 1, line drawings were derived for the

average shape (a), the high shape (b) from the 16 faces with highest attractiveness ratings and the caricatured high + 50% shape (c). Changes in feature positions from average to high + 50% shape are included in c. Composite face images with average (d), high (e) and high + 50% (f) shapes were obtained using the same methods as for Fig. 1.

average face shapes would exert a directional selection pressure away from the population mean. Selection pressures relating to other facial functions (for example, ingestion, respiration) may well work in different directions, limiting any evolutionary effects of preferences on face shapes. □

Received 7 October 1993; accepted 10 January 1994.

- Galton, F. *J. Nature* **18**, 97–100 (1878).
- Galton, F. *J. Anthropol. Inst. Gr. Br. Ir.* **8**, 132–142 (1878).
- Jastrow, J. *Science* **6**, 165–168 (1885).
- Stoddard, J. T. *Science* **8**, 89–91 (1886).
- Langlois, J. H. & Roggman, L. A. *Psychol. Sci.* **1**, 115–121 (1990).
- Light, L. L., Hollander, S. & Kayra-Stuart, F. *Personal. social Psychol. Bull.* **7**, 269–276 (1981).
- Cunningham, M. R. *J. pers. soc. Psychol.* **50**, 925–935 (1986).
- Cunningham, M. R., Barbee, A. P. & Pike, C. L. *J. pers. soc. Psychol.* **59**, 61–72 (1990).
- Alley, T. R. & Cunningham, M. R. *Psychol. Sci.* **2**, 123–125 (1991).
- ten Cate, C. & Bateson, P. *Evolution* **42**, 1355–1358 (1991).
- Andersson, M. *Nature* **299**, 818–820 (1982).
- Kirkpatrick, M. & Ryan, M. J. *Nature* **350**, 33–38 (1991).
- Møller, A. P. *Nature* **357**, 238–240 (1992).
- Enquist, M. & Arak, A. *Nature* **361**, 446–448 (1993).
- Cross, J. F. & Cross, J. *Dev. Psychol.* **5**, 433–439 (1971).
- Bernstein, I. H., Lin, T. & McLellan, P. *Percept. Psychophys.* **32**, 495–503 (1982).
- Symons, D. *The Evolution of Human Sexuality* (Oxford Univ. Press, New York, 1979).
- Brennan, S. E. *Leonardo* **18**, 170–178 (1985).
- Benson, P. J. & Perrett, D. I. *Image vis. Comput.* **9**, 123–129 (1991).
- Benson, P. J. & Perrett, D. I. *Perception* **22**, 257–262 (1993).

ACKNOWLEDGEMENTS. This work was supported by Unilever Research and the ESRC. We thank D. Jack, A. Mistlin, S. Marshall, P. Benson, A. Kinghorn, D. Rowland and M. Burt for help in processing faces, L. Pettigrew, R. Wells, T. Watson, L. Mackie, M. Booth and M. Oram for help in collecting and processing data, and M. Ritchie, G. Graves, R. Johnston, R. Byrne, R. Barton, N. Humphrey and P. Bateson for comments on drafts.

Simultaneous LTP of non-NMDA- and LTD of NMDA-receptor-mediated responses in the nucleus accumbens

Samuel B. Kombian & Robert C. Malenka*

Departments of Psychiatry and Physiology, LPPI, Box 0984, University of California, San Francisco, California 94143-0984, USA

THE nucleus accumbens (NA), a ventral extension of the striatum, plays a role in several complex behaviour patterns¹ and also is a major site of action of drugs of abuse such as cocaine^{1–3}. Intrinsic NA cells are predominantly quiescent^{4,5} and their activity depends on excitatory input from cortical and subcortical limbic afferents^{6,7}. Here we examine the mechanisms of synaptic plasticity at the synapse between prefrontal cortical afferents and cells in the core region of the NA^{8–10}. Manipulations that induce a Ca²⁺-dependent long-term potentiation (LTP) of non-NMDA (N-methyl-D-aspartate)-receptor-mediated responses also produce a simultaneous long-term depression (LTD) of NMDA-receptor-mediated responses. These results indicate that in a single cell the same change in postsynaptic Ca²⁺ concentration can have opposite effects on non-NMDA- and NMDA-receptor-mediated synaptic responses. This may be particularly important in the NA, where NMDA receptors are critical for mediating the behavioural actions of drugs of abuse^{11–13}.

Prefrontal cortical afferents make direct monosynaptic connections with cells in the NA core region and elicit an excitatory postsynaptic potential (e.p.s.p.) mediated mainly by non-NMDA receptors (non-NMDARs)^{9,10,14}. Tetanic stimulation of this input (100 Hz, 1 s) reliably induced LTP of both extracellular

field e.p.s.ps (62 ± 20%; mean ± s.e.; n = 14; Fig. 1a) and e.p.s.ps recorded using whole-cell recording techniques (68 ± 9%; n = 17; Fig. 1b). LTP induction was blocked reversibly by application of the NMDA receptor (NMDAR) antagonist D(-)-2-amino-5-phosphonovalerate¹⁵ (D-APV, 25 μM) (field e.p.s.p., Fig. 1a; n = 4; whole-cell e.p.s.p., n = 6; results not shown). LTP could also be induced by pairing moderate postsynaptic depolarization with low-frequency afferent stimulation (2 Hz for 2 min) (60 ± 6%; n = 7; Fig. 1c). To determine whether LTP in NA cells requires a rise in postsynaptic calcium ion concentration ([Ca²⁺]_i), electrodes were filled with the calcium chelator 1,2-bis(2-aminophenoxy)ethane-N, N, N', N'-tetraacetic acid (BAPTA, 10 mM). Tetanic stimulation did not elicit LTP in cells that had been loaded with BAPTA (Fig. 1d), whereas the increase in the simultaneously recorded field e.p.s.ps confirmed that the tetanic stimulation did induce LTP in surrounding cells (n = 4). Thus, like several forms of LTP found in hippocampus¹⁶ and in cortex¹⁷, LTP in the NA requires an NMDAR-dependent increase in [Ca²⁺]_i.

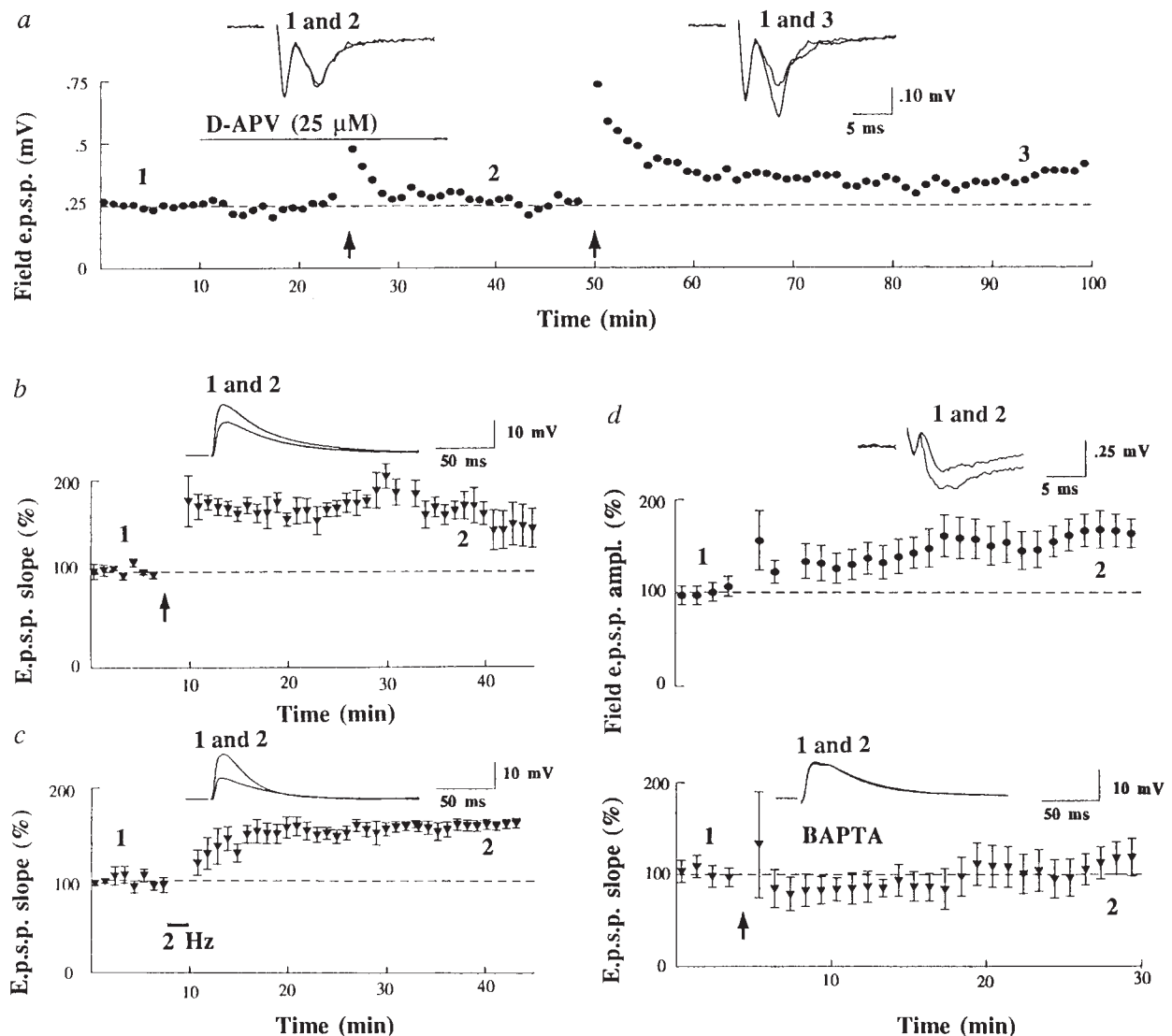
Because NMDA receptors contribute to synaptic transmission at these synapses^{4,5,8}, we examined whether tetanic stimulation also induced LTP of NMDAR-mediated synaptic responses. Surprisingly, tetanic stimulation caused a long-term depression (LTD) of the NMDAR-mediated e.p.s.p. (−41 ± 6%; n = 4; Fig. 2a, b) which was recorded in the presence of the selective non-NMDAR antagonist 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX; 10 μM) and 0.1 mM extracellular Mg²⁺. To test whether this change from LTP to LTD was due to the removal of Mg²⁺ from the perfusing medium, we isolated the NMDAR-mediated e.p.s.p. in normal medium (1.2 mM Mg²⁺) by depolarizing the cell in the presence of CNQX. Again, tetanic stimulation induced LTD of the NMDAR-mediated e.p.s.p. (−37 ± 10%; n = 6; Fig. 2c, d). Following its induction, LTD routinely lasted for the duration of the recording (50–70 min).

The induction of LTD versus LTP may depend on the postsynaptic [Ca²⁺]_i, with lower [Ca²⁺]_i causing LTD and higher [Ca²⁺]_i being required for LTP^{18,19}. Because non-NMDARs were blocked in the previous experiments, tetanic stimulation may not have caused as large an increase in postsynaptic [Ca²⁺]_i as that elicited in the absence of CNQX. To test this possibility, we examined the effects of pairing low-frequency afferent stimulation with direct postsynaptic depolarization and found that this protocol also induced LTD of the NMDAR-mediated e.p.s.p. (−48 ± 12%; n = 5) (Fig. 3a, b). LTD was blocked by loading cells with BAPTA (10 mM; n = 7; Fig. 3c) and also required NMDAR activation, as application of the tetanus in the presence of D-APV (25 μM) resulted in the complete recovery of the NMDAR-mediated response following D-APV washout (n = 3; results not shown).

Although the same induction protocols produced opposite effects on the two components of the e.p.s.p., an important question is whether LTP and LTD can be generated simultaneously in a single cell. To monitor non-NMDAR- and NMDAR-mediated synaptic components simultaneously, we took advantage of their different voltage dependence and kinetics^{20,21}. Voltage-clamping cells at −80 mV permitted measurement of non-NMDAR-mediated fast inward excitatory postsynaptic currents (e.p.s.cs), and holding the membrane potential at +20 to +30 mV allowed measurement (45 ms after stimulation) of predominantly NMDAR-mediated outward e.p.s.cs (Fig. 4A). After baseline e.p.s.cs were obtained at −80 mV, the cell was

FIG. 1 The induction of LTP of the non-NMDAR-mediated e.p.s.p. a, Bath application of D-APV (25 μM) blocks the induction of LTP in field e.p.s.ps recorded in the core region of the nucleus accumbens. After 10–15 min wash-out of D-APV, LTP could be induced (n = 4). In this and other figures, 'up' arrows represent times at which tetanic stimulation was applied. Insets show superimposed field e.p.s.ps taken at the indicated times. The first negative peak in the field e.p.s.p reflects direct

* To whom correspondence should be addressed.



activation of neuronal elements; the second peak is due to synaptic activation of intrinsic neurons¹⁵. *b*, Summary of experiments in which e.p.s.ps were recorded using whole-cell recording techniques and LTP was elicited by tetanic stimulation ($n=17$). In this and most other figures, the first and second post-tetanic responses have been deleted because they frequently were large and out of the range of the graph. Inset shows superimposed e.p.s.ps taken at the indicated times. *c*, Summary of experiments ($n=7$) in which LTP was induced by pairing low-frequency afferent stimulation with postsynaptic depolarization. Inset shows superimposed e.p.s.ps taken at the indicated times. *d*, Loading cells with BAPTA (10 mM; $n=4$) blocked the induction of LTP (bottom), although the simultaneously recorded field e.p.s.p. exhibited LTP (top). Insets show sample e.p.s.ps taken at the indicated times. **METHODS.** Parasagittal nucleus accumbens slices (400 μ m thick) were prepared from Sprague-Dawley rats (60–250 g) using standard techniques^{9,15}. Briefly, rats were anaesthetized with halothane, the brain was removed, and one hemisphere cut so that the lateral surface could be glued to a block and sliced with a vibratome in ice-cold artificial cerebrospinal fluid (CSF) which had been bubbled with 95% O₂ and 5% CO₂. The composition of the CSF was (in mM): NaCl, 126; KCl, 2.5; NaH₂PO₄, 1.2; MgCl₂, 1.2; CaCl₂, 2.4; NaHCO₃, 18; glucose, 11. After at least 1 h incubation at room temperature, a slice was transferred to a recording chamber and submerged beneath continuously flowing (2–3 ml per min) CSF (27–29 °C) to which 25 μ M picrotoxin had been added. Prelimbic cortical afferents were activated at 0.1 Hz (50–100 μ s, 10–40 V) by placing a bipolar stainless-steel microelectrode at the prefrontal–accumbens border, and recordings were made in the rostral–medial dorsal accumbens close to the anterior commissure. Field e.p.s.ps were recorded with glass microelectrodes (5–10 M Ω) filled with 3 M NaCl. Whole-cell current-clamp recording was performed using glass micropipettes (8–15 M Ω) filled with a solution containing (in mM):

potassium gluconate, 117.5; potassium methyl sulphate, 17.5; NaCl, 8; HEPES, 10; EGTA, 0.2; Mg-ATP, 2; GTP, 0.2, pH 7.2; osmolarity, 270–280 mosm. After addition of an aliquot of BAPTA (stock solution, 100 mM), pH and osmolarity of the recording solutions were readjusted. Blind whole-cell recordings were made²⁹ using an Axoclamp 2A amplifier (Axon Instruments). After break-in, action potential threshold, action potential amplitude, and the threshold stimulus strength for action potential generation were determined. Cells had resting potentials between –75 and –90 mV and spike amplitudes of at least 90 mV with no spontaneous activity. After collection of baseline responses (6–15 min), LTP or LTD was induced by tetanic stimulation given at the spike threshold stimulus strength (100 Hz for 1 s, given once or twice at 10-s intervals) or by pairing 2-Hz stimulation (2 min) with a 30–40 mV depolarization from the resting membrane potential. In current-clamp (bridge mode) experiments, the resting potential was always kept within 5 mV of the initial resting potential by current injection. Input resistance and bridge balance were constantly monitored by applying 20–50 pA (50–100 ms) current pulses 100–200 ms after the onset of the synaptic response. Cells were not included in data analysis if input resistance changed by more than 5–10%. Data were collected and analysed online (2–10 kHz sampling rate) using a 386-based personal computer programmed with Axobasic (Axon Instruments). E.p.s.p. initial slope (calculated using a least-squares regression) or amplitude (field e.p.s.ps, NMDAR-mediated e.p.s.ps) were used as a measure of synaptic efficacy. Summary graphs were constructed as described³⁰. In all figures, both data points and raw data traces are the average of six successive responses. Drugs were added to CSF from stock solutions just before application. CNQX and D-APV were from Cambridge Research Biochemicals, BAPTA was from Molecular Probes, and all other chemicals were from Sigma.

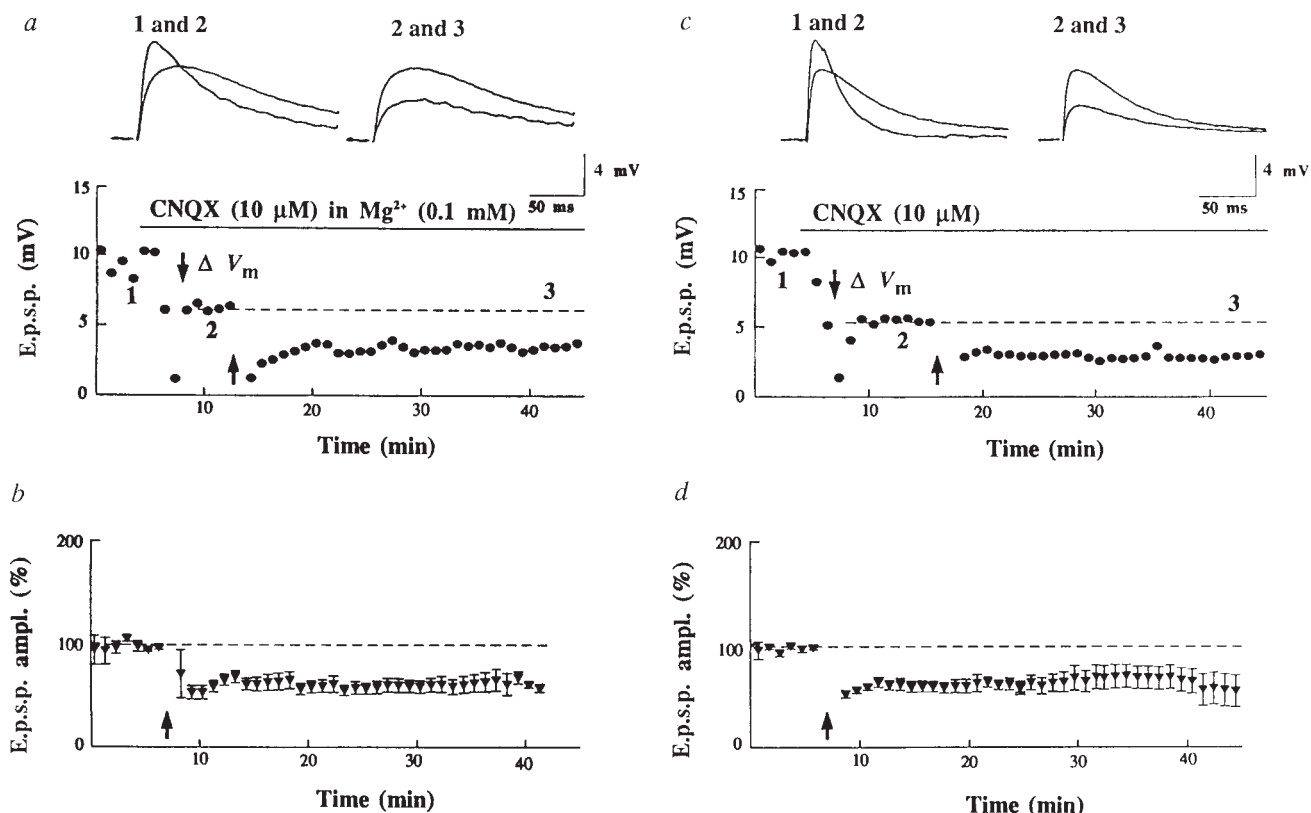


FIG. 2 Tetanic stimulation produces long-term depression of pharmacologically isolated NMDAR-mediated e.p.s.p.s. *a*, Tetanic stimulation (upward arrow) induced LTD of the NMDAR-mediated e.p.s.p. isolated by adding CNQX (10 μ M) to the perfusing medium and lowering Mg^{2+} (to 0.1 mM). The cell was depolarized (to -65 mV) at the point indicated by the downward arrow. Insets show superimposed responses taken at the indicated times. *b*, Summary of experiments ($n=4$) of the type

shown in *a*, *c*, Tetanic stimulation (upward arrow) induced LTD of the NMDAR-mediated e.p.s.p. recorded in normal CSF to which 10 μ M CNQX was added. The cell was depolarized (to -60 mV) at the point indicated by the downward arrow. Insets show superimposed synaptic responses taken at the indicated times. *d*, Summary of experiments ($n=5$) of the type shown in *c*.

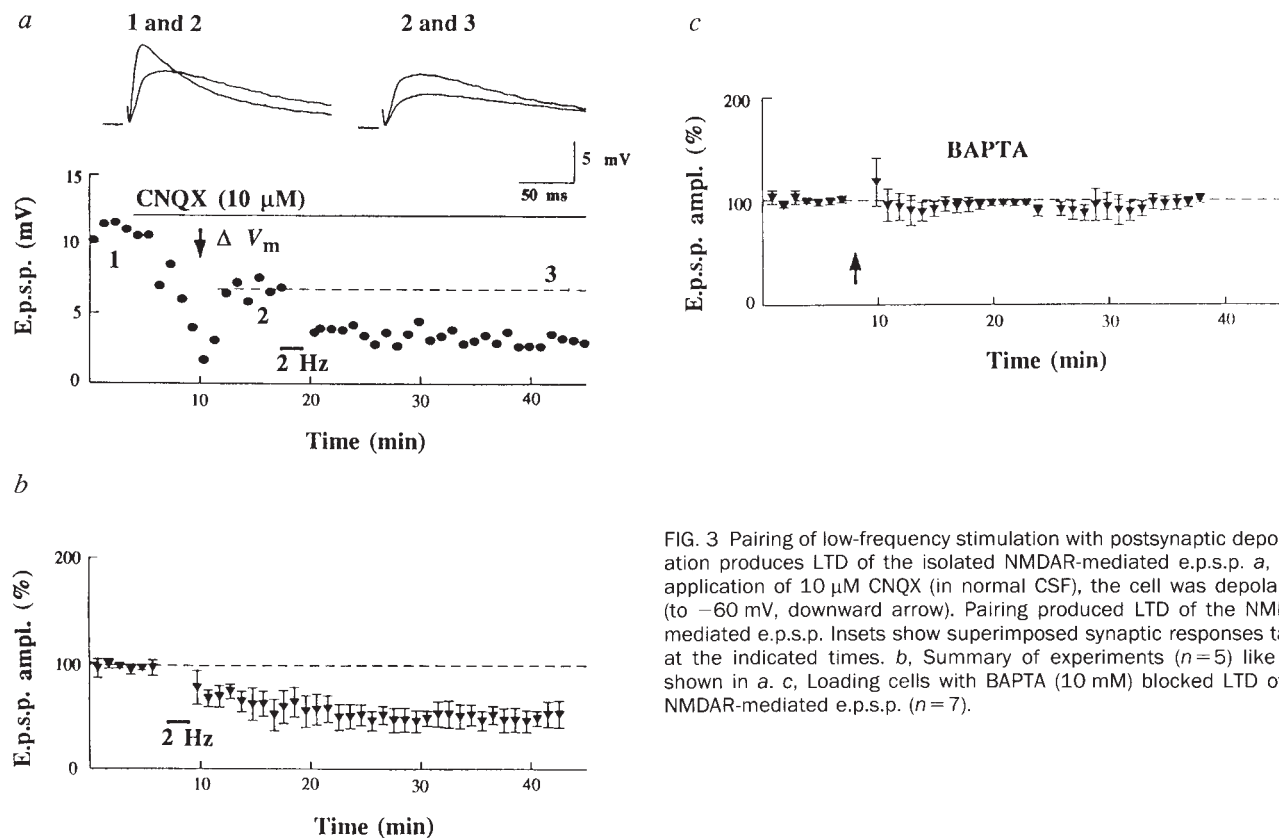


FIG. 3 Pairing of low-frequency stimulation with postsynaptic depolarization produces LTD of the isolated NMDAR-mediated e.p.s.p. *a*, After application of 10 μ M CNQX (in normal CSF), the cell was depolarized (to -60 mV, downward arrow). Pairing produced LTD of the NMDAR-mediated e.p.s.p. Insets show superimposed synaptic responses taken at the indicated times. *b*, Summary of experiments ($n=5$) like that shown in *a*. *c*, Loading cells with BAPTA (10 mM) blocked LTD of the NMDAR-mediated e.p.s.p. ($n=7$).

depolarized to +20–30 mV and the early and late components of the e.p.s.c. were measured during continued afferent stimulation²¹ (Fig. 4*B, C*). This resulted in a gradual decrease in the NMDAR-mediated outward e.p.s.c. ($-44 \pm 10\%$; $n=9$; Fig. 4*D*) and an increase in the predominantly non-NMDAR-mediated outward e.p.s.c. ($45 \pm 34\%$; Fig. 4*D*). Returning the cell to -80 mV confirmed the induction of LTP of the non-NMDAR-mediated e.p.s.c. ($81 \pm 17\%$; Fig. 4*B–D*). The decrease in the late component of the e.p.s.c. was not associated with any significant change in input resistance or series resistance (Fig. 4*B*) and was blocked by loading cells with BAPTA (10 mM; $n=2$).

LTD of the NMDAR-mediated e.p.s.c. reflects, at least in part, a true decrease in synaptic conductance (and not a shift in voltage-dependence²²) because it was observed at positive membrane potentials at which the Mg^{2+} -dependent block of the NMDAR channel is relieved^{23,24}. It also appears to be distinct from the homosynaptic LTD in hippocampus^{25,26} as prolonged (5–10 min) 1 Hz stimulation in the NA did not induce LTD of the NMDAR-mediated e.p.s.p. ($-5 \pm 10\%$; $n=4$), although the same prolonged stimulation did cause a depression of the non-NMDAR-mediated e.p.s.p. ($-45 \pm 6\%$; $n=6$; results not shown).

These results indicate that the same NMDAR-mediated

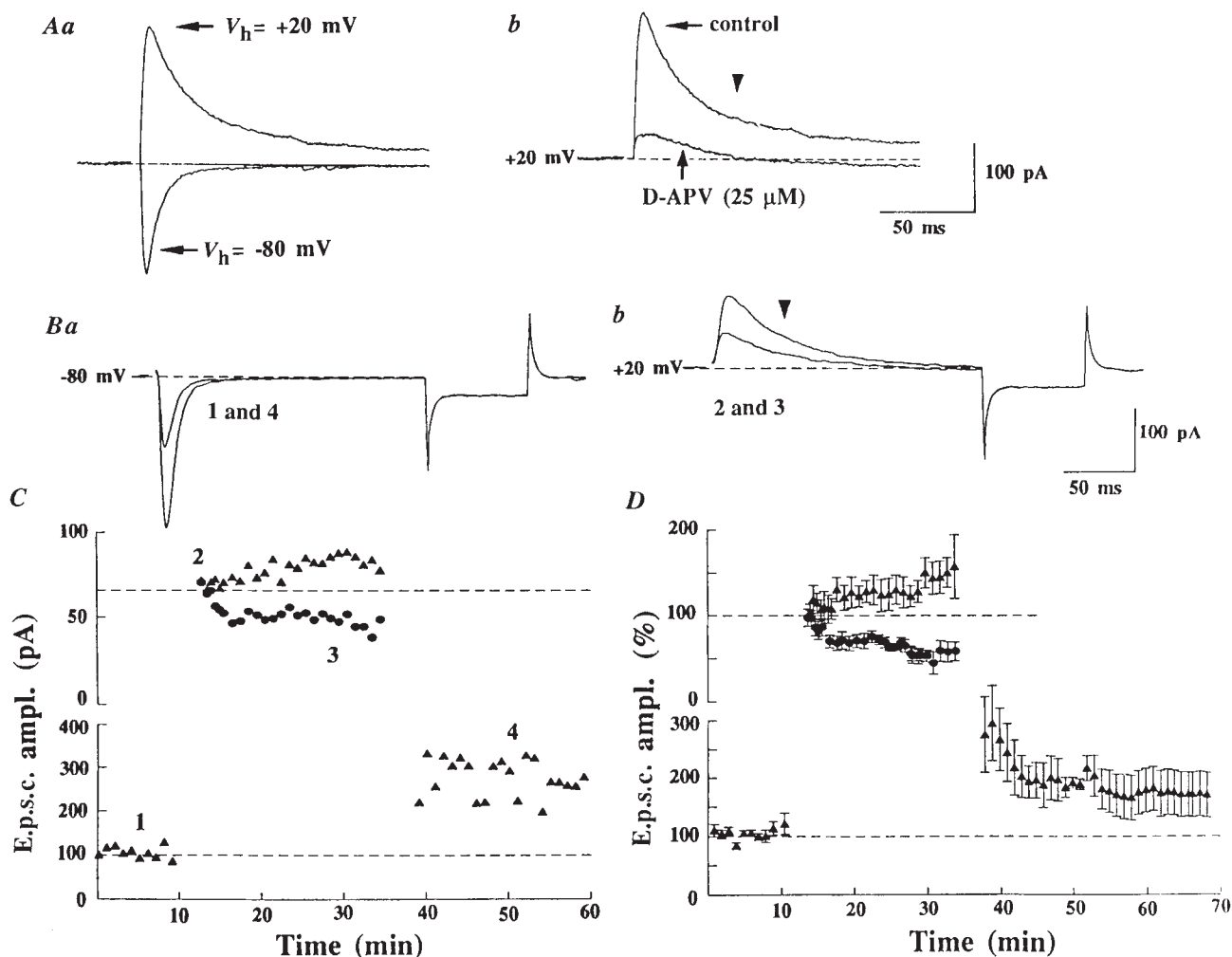


FIG. 4 LTP of the non-NMDAR-mediated e.p.s.c. and LTD of the NMDAR-mediated e.p.s.c. can be generated simultaneously in the same cell. *A, a*, Sample records (average of 3 successive responses) of e.p.s.c.s recorded at -80 mV and $+20$ mV holding potential (V_h). *A, b*, At $+20$ mV, D-APV ($25 \mu\text{M}$) reduced the peak e.p.s.c. by $64 \pm 15\%$ ($n=6$; mean $\pm$ s.d.). The arrowhead marks 45 ms after afferent stimulation, a time at which the NMDAR-mediated current comprises more than 95% of the synaptic response. *B*, Superimposed e.p.s.c.s taken at the indicated times in *C*: *a*, 1 and 4; *b*, 2 and 3. *C*, Bottom, peak e.p.s.c. amplitude at -80 mV, and top, e.p.s.c. amplitude at $+20$ mV measured at two time points (circles, 45 ms; triangles, 2 ms). *D*, Summary of experiments ($n=9$) like that shown in *C*. Six of these cells were held at $+20$ – 30 mV for 15 min while three cells were held at $+20$ – 30 mV for 24–25 min. For the bottom graph, values obtained after the return of the holding potential to -80 mV were offset by 10 min for the cells held at $+20$ – 30 mV for 15 min.

METHODS. Cells were voltage-clamped in the continuous single-electrode voltage-clamp mode. The pipette recording solution was the same as that used for the current-clamp experiments, except that K^+ was replaced by an equal amount of Cs^+ . Input resistance was constantly

monitored by applying a 10-mV (100 ms) hyperpolarizing step 100–200 ms after synaptic stimulation. It did not change during the course of the experiment (140 ± 14 (s.d.) $M\Omega$ before depolarization; 138 ± 15 $M\Omega$ following depolarization; $n=9$). Series/access resistance (30 ± 8 $M\Omega$) was monitored continuously by examining the capacitive transient at the beginning of the hyperpolarizing step. For all cells included in the data analysis, it did not vary by more than 5–10% during the course of the experiment. When the holding potential was changed, it was often necessary to wait 2–5 min before the holding current stabilized²¹. Once the holding current stabilized it also did not vary by more than 5–10%. The amplitude of the current within the initial 2–4 ms of the reversed e.p.s.c. at $+20$ to $+30$ mV was taken as a measure of the non-NMDA component of the e.p.s.c. (compare ref. 21). This early measurement, however, may be contaminated by NMDAR-mediated currents as D-APV often reduced the early component of the e.p.s.c. at positive holding potentials (*A*). Later components (>40 ms) of the reversed e.p.s.c. were completely blocked by D-APV and the amplitude of the NMDAR-mediated e.p.s.c. was measured at 45 ms (arrowheads in *A, b* and *B, b*) after synaptic stimulation.

increase in postsynaptic $[Ca^{2+}]_i$ can cause opposite effects on non-NMDAR- and NMDAR-mediated synaptic transmission at the synapses between prelimbic cortical afferents and NA core neurons. This is in marked contrast to studies of hippocampal synapses, where LTP of the non-NMDAR-mediated e.p.s.p. is accompanied by minimal or significant potentiation of NMDAR-mediated e.p.s.ps¹⁶. A plausible hypothesis for the difference in synaptic plasticity between NA and other brain regions is that NA cells are predominantly GABAergic and thus may express a different population of glutamate-receptor sub-

units from pyramidal cells²⁷. For example, *in situ* hybridization indicates that the striatum and NA contain levels of messenger RNA for NMDAR subunits that are different from those in the hippocampus and cortex²⁸. LTD of NMDAR-mediated synaptic transmission in the NA may be a mechanism for regulating or limiting the magnitude of change in the predominant non-NMDAR-mediated component of synaptic transmission. It may also be particularly important for limiting untoward psychomotor stimulatory effects that are mediated by NMDARs in the NA^{11,12}. □

Received 11 June 1993; accepted 6 January 1994.

1. Swerdlow, N. R. & Koob, G. F. *Behav. Brain Sci.* **10**, 197–245 (1987).
2. Nestler, E. J. *J. Neurosci.* **12**, 2439–2450 (1992).
3. Koob, G. F. & Bloom, F. E. *Science* **242**, 715–723 (1988).
4. Uchimura, N., Higashi, H. & Nishi, S. *J. Neurophysiol.* **61**, 769–779 (1989).
5. Pennartz, C. M. A., Boeijinga, P. H., Kitai, S. T. & Lopez da Silva, F. H. *Expl. Brain Res.* **86**, 190–198 (1991).
6. Christie, M. J., Summers, R. J., Stephenson, J. A., Cook, C. J. & Beart, P. M. *Neuroscience* **22**, 425–439 (1987).
7. Philipson, O. T. & Griffiths, A. C. *Neuroscience* **16**, 275–296 (1985).
8. Uchimura, N. & North, R. A. *J. Pharmac. exp. Ther.* **258**, 663–668 (1991).
9. O'Donnell, P. & Grace, A. A. *Synapse* **13**, 135–160 (1993).
10. Pennartz, C. M. A., Dölleman-Van der Weel, M. J. & Lopes da Silva, F. H. *Neurosci. Lett.* **136**, 109–112 (1992).
11. Kelly, A. E. & Throne, L. C. *Brain Res. Bull.* **29**, 247–254 (1992).
12. Pulvirenti, L., Swerdlow, N. R. & Koob, G. F. *Pharmac. Biochem. Behav.* **40**, 841–845 (1991).
13. Pulvirenti, L., Maldonado-Lopez, R. & Koob, G. F. *Brain Res.* **594**, 327–330 (1992).
14. Chang, H. T. & Kitai, S. T. *Brain Res.* **366**, 392–396 (1986).
15. Pennartz, C. M. A., Ameerun, R. F., Groenewegen, H. J. & Lopez da Silva, F. H. *Eur. J. Neurosci.* **5**, 107–117 (1993).

16. Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
17. Bear, M. F. & Kirkwood, A. *Curr. Opin. Neurobiol.* **3**, 197–202 (1993).
18. Lisman, J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9574–9578 (1989).
19. Artola, A. & Singer, W. *Trends Neurosci.* **16**, 480–487 (1993).
20. Hestrin, S., Nicoll, R. A., Perkel, D. J. & Sah, P. *J. Physiol. Lond.* **422**, 203–225 (1990).
21. Perkel, D. J. & Nicoll, R. A. *J. Physiol. Lond.* **471**, 481–500 (1993).
22. Chen, L. & Mae Huang, L.-Y. *Nature* **356**, 521–523 (1992).
23. Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. *Nature* **309**, 261–263 (1984).
24. Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, Z. *Nature* **307**, 462–465 (1984).
25. Dudek, S. M. & Bear, M. F. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4363–4367 (1992).
26. Mulkey, R. M. & Malenka, R. C. *Neuron* **9**, 967–975 (1992).
27. McBain, C. J. & Dingledine, R. *J. Physiol., Lond.* **462**, 373–392 (1993).
28. Monyer, H. *et al. Science* **256**, 1217–1221 (1992).
29. Bianton, M., Lo Turco, J. & Kriegstein, A. *J. Neurosci. Meth.* **30**, 203–210 (1989).
30. Malenka, R. C. *et al. Nature* **340**, 554–557 (1989).

ACKNOWLEDGEMENTS. We thank W. Colmers, C. Herron, R. Mulkey, R. Nicoll, S. Nicola, D. Perkel and D. Spillane for discussion. Supported by the NIMH, the McKnight Foundation, the National Alliance for Research on Schizophrenia and Depression (R.C.M.) and by the Human Frontiers Science Program Organization (S.B.K.).

Severe sensory and sympathetic neuropathies in mice carrying a disrupted *Trk*/NGF receptor gene

Richard J. Smeyne*, Rüdiger Klein*†, Andreas Schnapp†, Linda K. Long*, Sherri Bryant*, Anne Lewin*, Sergio A. Lira* & Mariano Barbacid*

* Department of Molecular Biology, Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, New Jersey 08543-4000, USA

† Differentiation Programme, European Molecular Biology Laboratory, 69012 Heidelberg, Germany

NERVE growth factor (NGF) induces neurite outgrowth and promotes survival of embryonic sensory and sympathetic neurons in culture^{1,2}. *In vivo*, NGF decreases the extent of naturally occurring cell death in developing sympathetic ganglia and protects cholinergic neurons of the basal forebrain and caudatoputamen^{1–3}. NGF interacts with the low-affinity p75 receptor and with *Trk*, a receptor tyrosine kinase encoded by the *trk* proto-oncogene^{4,5}. To study the role of *Trk* *in vivo*, we have ablated the gene in embryonic stem cells by homologous recombination. Mice lacking *Trk* have severe sensory and sympathetic neuropathies and most die within one month of birth. They have extensive neuronal cell loss in trigeminal, sympathetic and dorsal root ganglia, as well as a decrease in the cholinergic basal forebrain projections to the hippocampus and cortex. These findings demonstrate that *Trk* is the primary mediator of the trophic actions of NGF *in vivo* and that this signalling pathway plays a crucial role in the development of both the peripheral and the central nervous systems.

The *trk* gene was inactivated by transfecting D3 embryonic stem (ES) cells with a targeting vector in which 51 amino acids from the *Trk* kinase catalytic domain have been replaced with a phosphoglycerate kinase (PGK)-*neo* cassette in the opposite

orientation. The targeting vector also contained a PGK-thymidine kinase cassette for negative selection of non-homologous recombinants (Fig. 1a). Of 1,050 resistant clones, four were true recombinants as determined by polymerase chain reaction and Southern blot analysis (Fig. 1b). One clone was used to generate heterozygous mutant mice, which grow normally and have no obvious defects. Crossing these mice yielded homozygous animals with a frequency of 18% ($n=659$), indicating that mice lacking functional *Trk* receptors can develop to birth.

To examine *Trk* expression, we did reverse transcription-polymerase chain reaction (RT-PCR) and RNase protection assays with RNA isolated from 10-day-old (P10) wild-type, heterozygous and homozygous mutant mice. Our analysis confirmed the absence of normal *trk* transcripts; RNase protection analysis using a probe extending beyond the deleted region also failed to identify any *trk*-derived transcripts, suggesting that the insertional mutation impairs either the synthesis or the stability of the mutant transcripts (data not shown).

The homozygous mutant mice appear normal at birth, but by P10 they are noticeably smaller than their littermates, and about half of them die by P20. So far, none has survived beyond 55 days. At P10 they display a variety of sensory defects. They fail to react to deep pinpricks in their whisker pads and rear paws, indicating defects in their trigeminal and peripheral sensory nervous systems. As in *trkB* null mice⁶, they do not orient in response to vibrissal stimulation. In addition, they remain on a hot plate at 60 °C for at least 10 s (at which time they are removed), whereas wild-type and heterozygous mice jump off within 1–2 s.

By P20, surviving animals have myotic pupils and slight ptosis; when their eyes are illuminated, the iris constricts without efficient redilation, indicating that the parasympathetic component of the oculomotor nerve is intact but the sympathetic innervation from the superior cervical ganglion (SCG) is defective. In contrast, they respond normally to sound (handclap) and can discriminate 1% acetic acid from water, indicating that facial (VII), acoustic (VIII) and glossopharyngeal (IX) nerves function correctly. Basic motor functions also appear to be normal.

After 3 to 4 weeks, mutant mice develop a distinct pathology that worsens with increasing age. Their fur becomes mottled and

FIG. 1 Targeting of *trk* tyrosine kinase sequences by homologous recombination in ES cells. **a**, Schematic diagram of the strategy used to target the *trk* locus. Vertical black boxes represent exons encoding the kinase domain. Vertical thin lines outline the *trk* genomic sequences incorporated into the targeting vector pLL58. pLL58 has a 1.0 kilobase (kb) deletion encompassing the entire intron between exons C6 and C7, the 3' 60 base pairs (bp) of exon C6 and 5' 93 bp of exon C7 (amino acids 710–760 of Trk; ref. 18). Putative crossovers between the targeting pLL58 DNA and the endogenous *trk* locus are indicated as crossed dashed lines. The PGK-1/*neo* cassette¹⁹ is indicated by a hatched box labelled *neo*. The PGK-1/HSV thymidine kinase cassette²⁰ is indicated by an open box labelled *tk*. Arrows inside boxes indicate the direction of transcription. Cleavage sites for restriction endonucleases *Bgl*III (B), *Eco*RI (E), *Kpn*I (K), *Nco*I (Nc) and *Not*I (Nt) are indicated. The *Bgl*III site used for cloning pLL58 is indicated with square brackets. pLL58 was linearized with *Not*I before transfection. The region amplified by PCR to identify G418^R/Ganc^R ES cell clones carrying homologous recombination events is indicated by an open box. The 201-bp long DNA probe used to identify the diagnostic 3.0-kb *Bgl*III, 8.0-kb *Kpn*I and 10.3-kb *Eco*RI DNA fragments (arrows) is indicated by a small hatched box. **b**, Southern blot analysis of PCR-positive G418^R/Ganc^R ES cell clones. Genomic DNAs extracted from ES cell clones B151-37 and B170-316 and from untransfected D3 ES cells were digested with *Bgl*III, *Kpn*I or *Eco*RI, and analysed with the *trk* probe indicated in **a**. The wild-type *trk* allele generates 2.4-kb *Bgl*III, 4.7-kb *Kpn*I and 9.6-kb *Eco*RI fragments, whereas the targeted

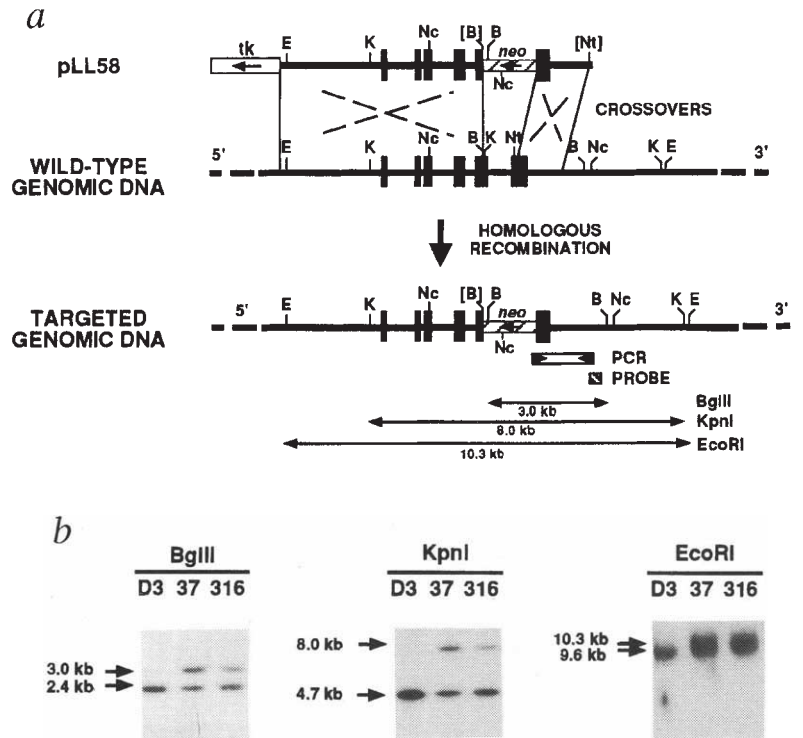
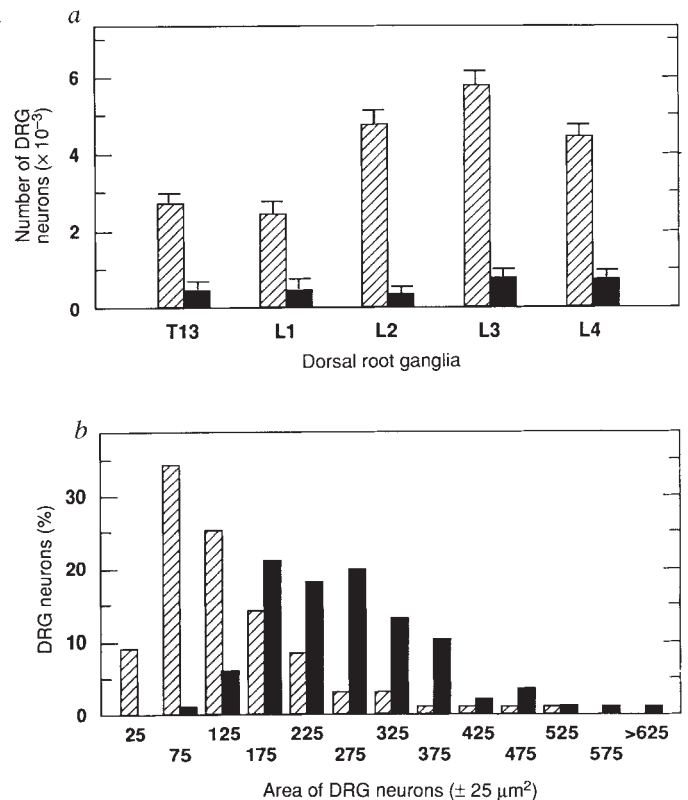


FIG. 2 Neuronal cell loss in the DRGs of *trk*^{TK} (–/–) mice. **a**, Cell counts of neurons in the thoracic (level 13) and lumbar (levels 1 to 4) DRGs of PO *trk*^{TK} (+/+) (hatched boxes) and *trk*^{TK} (–/–) (black boxes) mice. Counts of the DRG ganglia neurons were as previously described⁸. **b**, Histogram of cell soma area of DRG neurons in PO *trk*^{TK} (+/+) (hatched boxes) and *trk*^{TK} (–/–) (black boxes) mice. PO DRGs were embedded in paraffin and sectioned at 5 μ m. Area measurements were made by tracing the contours of 500 individual neuronal somata magnified at $\times 1,000$ onto a digitizing tablet. The area of the individual neurons was calculated using a morphometric program (SigmaScan, Jandel).



numerous scabs appear over the entire body; their paws are covered by ulcerations and digits are missing, probably due to self-mutilation; their eyes exhibit corneal opacities, indicating there may be a sensory defect affecting the blink response. Although mutant mice possess all of their internal organs, some organs are abnormal, but the cause of this has not been determined.

The *trk* gene is expressed in the trigeminal, dorsal root and sympathetic ganglia of the peripheral nervous system^{7,8}. Examination of PO mutant mice revealed that their trigeminal ganglia are considerably smaller than those of their wild-type siblings, and that the number of trigeminal neurons is reduced by 70–90% (5,383 $\pm$ 101 neurons in (–/–) mice ($n=3$) compared to 21,132 $\pm$ 567 in (+/+) mice ($n=4$); $P<0.001$). In contrast to *trkB*-deficient mice⁶, the loss of neurons appears to be distributed throughout the ganglion. No significant abnormalities were observed in other cranial ganglia examined. The dorsal root ganglia (DRG), however, have extensive neuronal loss; neuronal counts (in DRGs T13 and L1–4) revealed a 70–90% loss of neurons in mutant ($n=3$) compared to wild-type ($n=4$) animals. To determine whether a particular class (small, medium or large) of neurons was absent, we measured the soma size of over 500 mutant and wild-types DRG neurons. Our results (Fig. 2b) indi-

cate that small (50–200 μ m²) neurons, which are believed to be NGF-dependent⁹, are preferentially lost in mutant mice. The SCG of perinatal mutant mice also show a loss of neurons as well as many pyknotic nuclei which suggest an active process of cell degeneration. By P10, the SCG is severely shrunken (Fig. 3b), although the few remaining principal sympathetic neurons

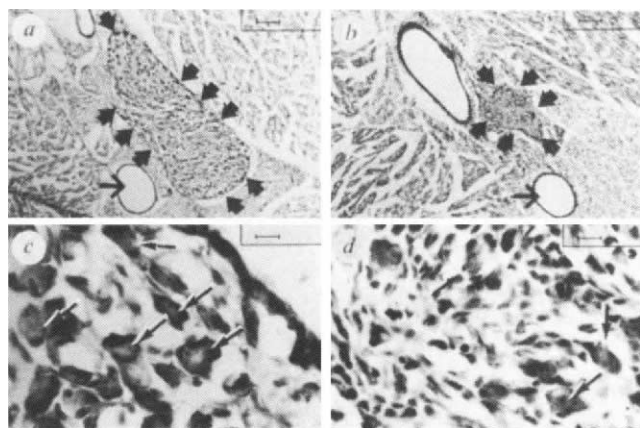


FIG. 3 Loss of sympathetic neurons in the SCG of *trk*^{TK} (-/-) mice. Nissl preparations of P10 mouse tissue showing the SCG (arrowheads) of *trk*^{TK} (+/+) (a) and *trk*^{TK} (-/-) (b) mice. The carotid artery is indicated by the arrow. The boxed areas in a and b are shown in higher magnification in c and d, respectively. The large principal sympathetic neurons are easily distinguished by their large size, dark staining cytoplasm, pale staining nucleus and distinct nucleoli (arrows). Scale bars a, c, 200 μ m; b, d, 20 μ m.

are normal in size and staining characteristics (Fig. 3c, d). The identity of these cells was confirmed by immunostaining for tyrosine hydroxylase (not shown). Similar levels of neuronal loss were observed in other sympathetic ganglia.

Expression of the *trk* gene in the central nervous system (CNS) is restricted to small subsets of cholinergic neurons in the caudatoputamen and the basal forebrain^{10,11}, two known targets for NGF³. To determine whether these neurons are affected by loss of Trk receptors, we stained sagittal sections of adult mutant and wild-type brains ($n=3$) for acetylcholinesterase (AChE)¹². Adult mutant mice display a substantial decrease in AChE staining in cholinergic fibres that project from the medial septum to the hippocampus (Fig. 4a, b) and from the nucleus basalis to the cerebral cortex (Fig. 4c, d). In contrast, staining is unaffected in the caudatoputamen (Fig. 4e, f) and in regions such as brainstem, which are not known to express *trk* (Fig. 4g, h).

The sensory and sympathetic neuropathies we observe in *trk*-deficient mice are probably due to the loss of neurons in the trigeminal, sympathetic and dorsal root ganglia, three well known targets for NGF^{1,13}. Earlier studies^{14,15} that used immunological methods to inhibit NGF activity during development demonstrated a role for NGF in the survival of sympathetic and sensory neurons, and the similarity between these defects and those we observe in *trk*-deficient mice strongly suggest that the *trk* receptor is required for most if not all of the trophic effects of NGF in the peripheral nervous system (PNS). The severity of the *trk*-deficient phenotype, as compared to the normal sympathetic nervous system and limited sensory deficiencies observed in mice lacking the low-affinity p75 NGF receptor¹⁶, strengthens the concept^{4,5} that Trk is the predominant receptor for NGF and that p75 plays an accessory rather than a direct role in mediating the functions of NGF *in vivo*.

The role of NGF in the CNS is less well documented, although NGF stimulates sprouting of basal cholinergic neurons *in vitro*, and administration of NGF *in vivo* protects septal cholinergic neurons from death following transection of their projections to the hippocampus, a presumed physiological source of NGF³. In adult *Trk*-deficient mice, cholinergic projections of the basal forebrain to the hippocampus and cerebral cortex are severely decreased. Preliminary studies (L. Kromer, personal communication) suggest that the number of neurons within the basal

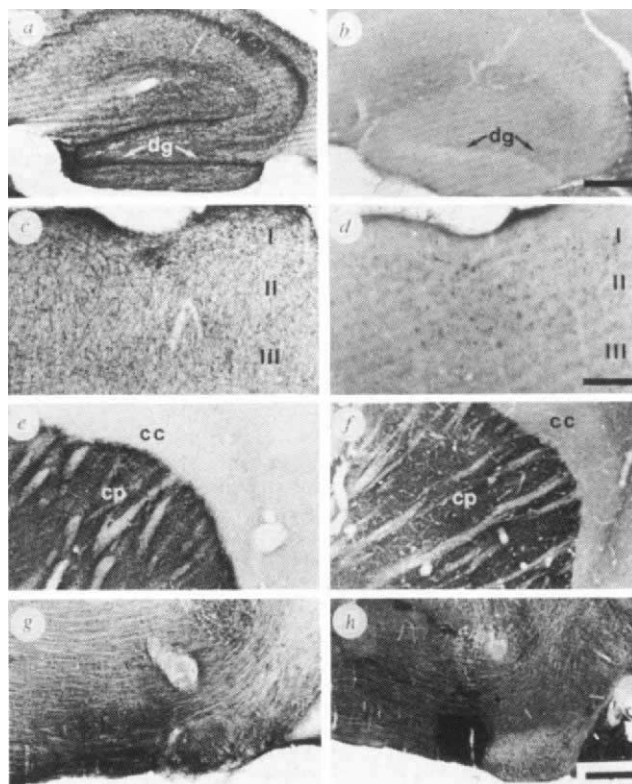


FIG. 4 Cholinergic projections in the brain of *trk*^{TK} (-/-) mice. Acetylcholinesterase (AChE)-positive fibres projecting to the hippocampus in adult *trk*^{TK} (+/+) (a) and *trk*^{TK} (-/-) (b) mice. The region surrounding the dentate gyrus is indicated by arrows. AChE-positive fibres projecting to the somatosensory cortex in adult *trk*^{TK} (+/+) (c) and *trk*^{TK} (-/-) (d) mice. Cortical layers I-III are indicated. AChE staining in the caudatoputamen (cp) of adult *trk*^{TK} (+/+) (e) and *trk*^{TK} (-/-) (f) mice. The corpus callosum (cc) which lies rostral and superior to the caudatoputamen, is indicated. AChE-positive fibres staining in the ventral brainstem of adult *trk*^{TK} (+/+) (g) and *trk*^{TK} (-/-) (h) mice. Histochemical detection of AChE activity was done in individual free-floating sections (20–40 μ m) as described¹². Scale bars a, b, 200 μ m; c, d, 50 μ m; e–h, 200 μ m.

forebrain complex is not significantly reduced in these mice, suggesting that the loss of cholinergic fibres may reflect a role for Trk either in fibre outgrowth or in the maintenance of the cholinergic phenotype. In any case, our results indicate that Trk is likely to be responsible for mediating the effects of NGF in the CNS. The fact that most known *trk*-expressing structures in both the PNS and CNS are affected by the *trk* -/- mutation contrasts with the phenotypes produced by disruption of many other genes, including the related neurotrophin receptors *trkB* and *trkC*^{6,17}, in which only a limited subset of the expressing cells are detectably affected. This suggests that loss of the NGF/Trk signalling pathway cannot be compensated for by other factors, and the *Trk* mutant mice described here will be an invaluable tool for determining the mechanisms by which neurotrophins contribute to the development and maintenance of the mammalian nervous system. □

Received 10 December 1993; accepted 28 February 1994.

1. Levi-Montalcini, R. *EMBO J.* **6**, 1145–1154 (1987).
2. Bradshaw, R. A., Blundell, R. L., Lapato, R., McDonald, N. Q. & Murray-Rust, J. *Trends biochem. Sci.* **18**, 48–52 (1993).
3. Dreyfus, C. F. *Trends pharmac. Sci.* **10**, 145–149 (1989).
4. Meakin, S. O. & Shooter, E. M. *Trends Neurosci.* **15**, 323–331 (1992).
5. Barbacid, M. *Oncogene* **8**, 2033–2042 (1993).
6. Klein, R. et al. *Cell* **75**, 113–122 (1993).
7. Martin-Zanca, D., Barbacid, M. & Parada, L. *Genes Dev.* **4**, 683–694 (1990).

8. Tessarollo, L. et al. *Development* **118**, 463–475 (1993).
9. Mu, X., Silos-Santiago, I., Carroll, S. L. & Snider, W. D. *J. Neurosci.* **13**, 4029–4041 (1993).
10. Holtzman, D. M. et al. *Neuron* **9**, 465–478 (1992).
11. Steininger, T. L., Wainer, B. H., Klein, R., Barbacid, M. & Palfrey, H. C. *Brain Res.* **612**, 330–335 (1993).
12. Butcher, L. L. & Woolf, N. J. in *Handbook of Chemical Neuroanatomy* Vol. II (eds Bjorkland, A., Hokfelt, T. & Kuhar, M. J.) (Elsevier, Amsterdam, 1984).
13. Davies, A. M., Lee, K.-F. & Jaenisch, R. *Neuron* **11**, 565–574 (1993).
14. Johnson, E. M. Jr, Gorin, P. D., Brandeis, L. D. & Pearson, J. *Science* **210**, 916–918 (1980).
15. Ruit, K. G., Elliott, J. L., Osborn, P. A., Yan, Q. & Snider, W. D. *Neuron* **8**, 573–587 (1992).

16. Lee, K. F. et al. *Cell* **69**, 737–749 (1992).
17. Klein, R. et al. *Nature* **368**, 249–251 (1994).
18. Martin-Zanca, D., Oskam, R., Mira, G., Copeland, T. & Barbacid, M. *Molec. cell. Biol.* **9**, 24–33 (1989).
19. McBurney, M. W. et al. *Nucleic Acids Res.* **20**, 5755–5761 (1991).
20. Mansour, S. L., Thomas, K. R. & Capecchi, M. R. *Nature* **336**, 348–352 (1988).

ACKNOWLEDGEMENTS. We thank A. Joyner, W. Wurst and A. B. Auerbach for introducing us to gene targeting, M. Swerdel, J. D. Wallace and P. Zalamea for technical assistance R. Brambilla for the pRB6 plasmid, and A. Fagan for the AChE histochemistry.

Disruption of the neurotrophin-3 receptor gene *trkC* eliminates Ia muscle afferents and results in abnormal movements

Rüdiger Klein^{*†}, Inmaculada Silos-Santiago[†], Richard J. Smeyne^{*}, Sergio A. Lira^{*}, Riccardo Brambilla[†], Sherri Bryant^{*}, Li Zhang[†], William D. Snider[†] & Mariano Barbacid^{*}

^{*} Department of Molecular Biology, Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, New Jersey 08543, USA

[†] Differentiation Programme, European Molecular Biology Laboratory, 69012 Heidelberg, Germany

[‡] Department of Neurology and Neurological Surgery, Washington University School of Medicine, St Louis, Missouri 63110, USA

THE *trkC* gene^{1,2} is expressed throughout the mammalian nervous system^{3–5} and encodes a series of tyrosine protein kinase isoforms that serve as receptors for neurotrophin-3 (NT3), a member of the nerve growth factor (NGF) family of neurotrophic factors^{2,6–8}. One of these isoforms, gp145^{trkC}/TrkC K1, mediates the trophic properties of NT3 in cultured cells^{2,6–8}. Here we show that homozygous mice defective for TrkC tyrosine protein kinase receptors lack Ia muscle afferent projections to spinal motor neurons and have fewer large myelinated axons in the dorsal root and posterior columns of the spinal cord. These mice display abnormal movements and postures, indicating that NT3/TrkC-dependent sensory neurons may play a primary role in proprioception, the sense of position and movement of the limbs.

Mice heterozygous for a targeted mutation in the tyrosine kinase domain of the *trkC* locus were generated as described in the legend to Fig. 1. These *trkC*^{TK} +/– mice grow normally and show no obvious anatomical or behavioural defects. Crossing these heterozygous mice produced homozygous –/– animals with a frequency of 21% (*n* = 285), indicating that mice lacking functional TrkC tyrosine kinase receptors can develop to birth.

To examine TrkC expression, we performed RNase protection assays with RNAs isolated from 10-day-old (P10) wild-type, heterozygous and homozygous mutant mice. A 313-nucleotide (nt) riboprobe encompassing the first (K1) and second (K2) kinase exons was completely protected by RNAs derived from wild-type and heterozygous mice but not by those isolated from the homozygous mutant mice (Fig. 1a). Instead, these transcripts protected a 130-nt fragment probably derived from the non-targeted K1 exon. To verify this, we used reverse transcription-polymerase chain reactions (RT-PCR) based on primers corresponding to sequences encompassing exons K1 and K2 as well as part of exon K3. These primers amplified the expected 419-base-pair (bp) DNA fragment from RNAs isolated from wild-type and heterozygous mice but only a 244-bp fragment from homozygous mutant mice (Fig. 1b). Both of these amplified complementary DNA fragments hybridized to an oligonucleotide corresponding to exon K1 sequences. But only the wild-type

419-bp DNA fragment hybridized to another oligonucleotide derived from exon K2, thus confirming the absence of exon K2 sequences in transcripts from homozygous mutant mice (Fig. 1b). Nucleotide sequence analysis of the mutant 244-bp DNA fragment amplified from homozygote-derived RNA revealed a precise splicing from exon K1 to exon K3 (Fig. 1c). The putative proteins encoded by these abnormal transcripts lack 58 amino-acid residues within their kinase domain and have no catalytic activity. Moreover, abnormal splicing introduces an extra nucleotide (G) to the targeted transcripts, completely changing the last 95 residues. The predicted structure of the targeted *trkC* transcripts and their putative products are shown in Fig. 1d.

The homozygous mutant mice appear normal at birth, but by P4 they are reduced in size and most of them die by P21. A small number, however, survive for four months or longer. By P4 it is apparent that these mice do not move normally, and by P10 they develop abnormal, ataxic movements when walking; even at rest, the limbs are held in abnormal positions in relation to the trunk. When placed on a rotating dowel, the homozygous mutant mice cannot maintain an upright posture and immediately tumble off. In addition, these mutant mice draw their limbs in towards their bodies when lifted by their tails, in contrast to the fully extended posture observed with their wild-type and heterozygous littermates. The mutant mice do, however, take nourishment and respond to painful stimuli in their whisker pad, two functions that are defective in mice lacking TrkB receptors⁹. Similarly, the *trkC* mutant mice respond to pinching of the tail

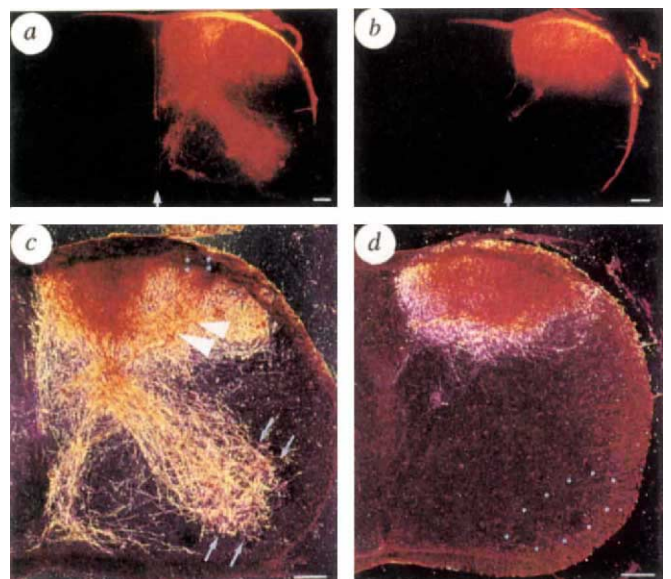
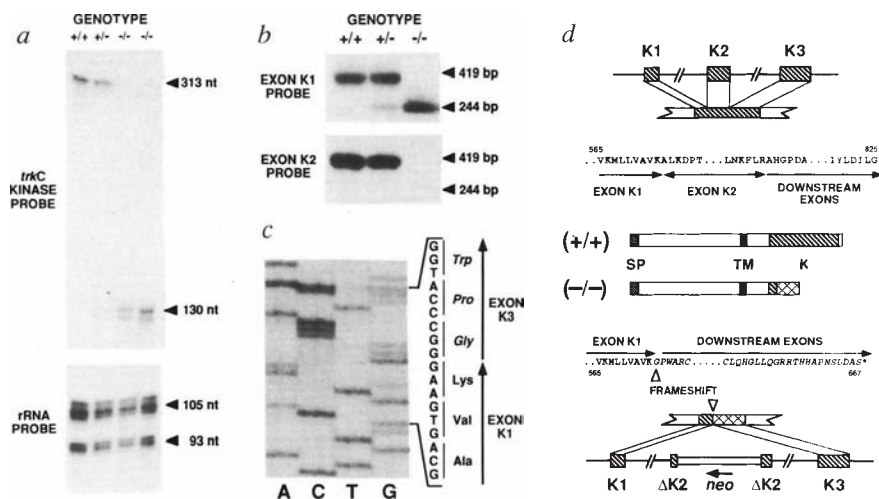


FIG. 2 Absence of Ia muscle afferent projections in *trkC*^{TK} –/– mice. Dil staining of the C5 dorsal root afferent projection to the spinal cord in PO +/+ (a, c) and –/– (b, d) mice. Small arrows in a and b denote the midline of the spinal cord. Large arrows in c indicate Ia muscle afferents in +/+ mice. Cutaneous (asterisks) and mechanoreceptive (arrowheads) afferents are also indicated. Scale bars, 100 μm.

FIG. 1 Generation of *trkC*^{TK} mutant mice. Genomic *trkC* sequences encoding the tyrosine kinase catalytic domain of gp145^{trkC} (amino-acid residues 573–630²) were used to generate a targeting vector, pFRK100, by replacing 21 bp of the kinase second exon (K2) with a phosphoglycerate kinase 1 (PGK1-*neo*) cassette²⁰. A thymidine kinase (tk) cassette used for negative selection of cells carrying nonhomologous recombinations²¹ was placed at the 5' end of these genomic sequences. Both *neo* and tk cassettes were inserted in the opposite transcriptional orientation as the *trkC* gene. ES cells (D3 clone) were transfected with pFRK100 DNA and screened with genomic sequences for homologous recombination events as previously described^{9,22}. Of a total of 580 double resistant ES cell clones screened, two (B149-339 and B149-361) were found to be true recombinants. Targeted ES cell clones were injected into C57Bl/6 blastocysts and transferred into the uteri of pseudopregnant CD1 recipient mothers^{9,22}. Chimaeras derived from the B149-361 clone transmitted the targeted allele when bred to C57Bl/6 mice. Crossing these *trkC*^{TK} $\pm$ heterozygous mice produces *trkC*^{TK} $\pm$ animals with a frequency of 21% ($n=285$). Molecular analysis of *trkC* transcripts in *trkC*^{TK} $\pm$ mice: **a**, RNase protection assay. Total RNAs were isolated from heads of P10 $\pm$ and $\pm$ and $\pm$ mice using the guanidine-thiocyanate method²³. RNase protection analysis was performed as described²⁴. Briefly, a 425-bp *Nco*I DNA fragment from a mouse *trkC* cDNA clone (nucleotides 1,487–1,929 in pFL19, see ref. 2) was subcloned into *Nco*I-digested pGEM-5Zf(-) (Promega). The resulting plasmid, pAS1, was linearized at an internal *Nde*I site and submitted to *in vitro* transcription with T7 polymerase. The resulting [³²P]-labelled 313-nt riboprobe was hybridized to 20 μ g total RNA, digested with RNases A and T1, phenol-extracted and precipitated. A [³²P]-rRNA probe was also used to control for the amount of loaded RNA. The protected [³²P]-RNA fragments were separated in an 8% polyacrylamide gel and visualized by autoradiography (7 days (*trkC* probe) or 2 h (rRNA probe)). The size of the protected RNA fragments (indicated by arrows) was determined by co-electrophoresis of [³²P]-labelled *Hpa*II-digested pBR322 DNA standards. **b**, RT-PCR assay. Total RNAs (50 μ g) from $\pm$, $\pm$ and $\pm$ mice were reverse-transcribed into cDNA using oligo(dT) primers and SuperScript Reverse Transcriptase (BRL) and amplified using primers specific for the tyrosine kinase domain of *trkC*.



The resulting PCR-amplified DNAs were submitted to Southern blot analysis using as probes oligonucleotides derived from exon K1 or exon K2 sequences. The size of the hybridized DNA fragments (indicated by arrows) was determined by co-electrophoresis of a 1-kb DNA ladder (BRL). **c**, Nucleotide sequence analysis of the 244-bp DNA fragment generated by RT-PCR of $\pm$ mouse RNA. Nucleotides derived from exons K1 and K3 are indicated along with the predicted amino-acid sequence. Note that this splice event generates a frameshift mutation that leads to an aberrant amino-acid sequence (italics) in exon K3. **d**, Schematic representation of the results depicted in **a**, **b** and **c**. Symbols include the PGK1-*neo* cassette (*neo*, stippled box) and the signal peptide (SP, grey boxes), transmembrane region (TM, black boxes), kinase exons (K1–K3, hatched boxes) and tyrosine kinase domain (K, hatched box) of the wild-type TrkC tyrosine kinase receptors. In *trkC*^{TK} $\pm$ mice (bottom half), the *neo*-disrupted K2 exon (Δ K2) is spliced out leading to an aberrant transcript that contains a frameshift mutation (open triangles) between exons K1 and K3. This mutation results in the synthesis of abnormal TrkC receptors of 667 amino acids (against 825 residues in gp145^{trkC}/TrkC K1) whose 95 C-terminal residues (cross-hatched box) bear no resemblance to any of the sequences present in wild-type TrkC receptors^{2,6–8}. Aberrant amino-acid residues (573–667) are indicated in italics.

and foot, unlike mice targeted at the related *trk* proto-oncogene locus¹⁰. These observations demonstrate distinct roles for each member of the neurotrophin receptor family in the mammalian nervous system.

The behavioural phenotype of the homozygous mutant mice suggest a defect in proprioception. To determine the cause of this defect, we examined dorsal root ganglia (DRG) which contain the primary afferent neurons that subserve proprioception¹¹. Lumbar DRGs of newborn wild-type mice contained $18,678 \pm 881$ neurons ($n=6$), whereas mutant $\pm/\pm$ littermates contained $15,095 \pm 1,009$ neurons ($n=6$): $t=2.443$; $P<0.05$. The observed 19% neuron loss in the homozygous mutant mice corresponds well with the subpopulation of DRG neurons known to express *trkC* transcripts^{12,13}.

To determine whether a particular class of dorsal root afferents was absent in the mutant mice, we stained cervical and thoracic DRGs with the lipid-soluble tracer 1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate (DiI) and visualized the lamina-specific projections of different classes of DRG neurons in the spinal cord. Three major groups of DRG afferents can be identified, all of which are well developed by birth¹¹. At this time, fascicles of axons projecting to the motor pools are seen in wild-type mice (Fig. 2a and c, long arrows). These projections are derived from DRG neurons that innervate primary endings of muscle spindles in the periphery, give off Ia axon¹⁴ collaterals to motor pools and project to the dorsal column nuclei. A second group of afferent projections enter the

spinal cord medially and project laterally to laminae III and IV in the dorsal horn (Fig. 2a and c, arrowheads). These projections presumably innervate low-threshold mechanoreceptors in the periphery¹⁴. The third group of DRG projections, the cutaneous afferents which subserve nociception and thermoception, densely innervate laminae I and II (Fig. 2a and c, asterisks).

Group Ia muscle afferents were absent in each section of cervical or thoracic spinal cords of eight mutant mice examined (Fig. 2b and d). Group II afferents which supply the secondary endings of muscle spindles were also apparently absent in mutant animals, but these fibres are much less common than Ia afferents in rodents¹⁴ and we did not systematically assess their presence in normal mice. In contrast, the dorsal horn innervation appears grossly normal in the mutant mice, indicating the presence of at least some classes of nociceptors and low-threshold mechanoreceptors (Fig. 2b and d). There may nevertheless be loss of some populations of afferents which normally project to the superficial or deep dorsal horn.

DRG neurons projecting to the motor pools normally have large axons that become heavily myelinated by the fourth postnatal week¹¹. We observed a marked depletion of these myelinated axons in adult (P30) mutant animals compared with their littermates ($2,176 \pm 194$ myelinated axons in wild-type mice ($n=3$) against $1,150 \pm 85$ myelinated axons in homozygous mutant mice ($n=4$): $t=9.61$; $P<0.0001$) (Fig. 3a and c). Because myelinated axons represent about 40% of all dorsal root axons, the observed axonal loss corresponds well with the 19% loss of neu-

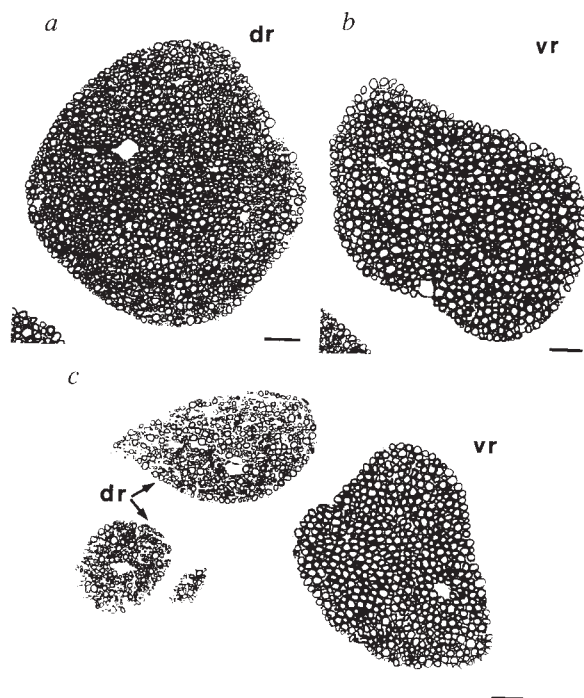


FIG. 3 Reduced numbers of large myelinated axons in the dorsal root and ventral roots of *trkC*^{TK} $-/-$ mice. Semithin sections of lumbar dorsal (dr) and ventral (vr) roots of adult (P30) $+/+$ (a, b) and $-/-$ (c) mice were stained for myelin. Scale bars, 25 μ m.

rons observed in the DRGs of newborn animals. But it remains possible that additional neurons are lost between birth and P30 in the mutant mice.

Some Ia afferent axons ascend to the nucleus gracilis and nucleus cuneatus through the dorsal columns¹⁵. To determine whether these ascending projections are affected in the mutant mice, we measured the cross-sectional area of the posterior columns in newborn mice. Whereas the dorsal columns of wild-type mice ($n=38$) measure $99,861 \pm 22,944 \mu\text{m}^2$, those of the homozygous mutant mice ($n=66$) only measure $49,369 \pm 11,692 \mu\text{m}^2$ ($t=14.68$; $P<0.0001$). The large myelinated group Ia axons which ascend in the dorsal columns and extend collaterals to motor pools and spinocerebellar neurons are therefore either absent or drastically depleted in the mutant mice. This defect may be largely responsible for the behavioural deficits observed in these mice.

Loss of other neuronal populations may also contribute to the mutant phenotype. For instance, the ventral roots of the spinal cord containing the motor neuron axons show moderate loss of fibres in the mutant mice (941 ± 12 myelinated axons in wild-type mice ($n=3$) against 643 ± 67 myelinated axons in homozygous mutant animals ($n=4$): $t=4.01$; $P<0.01$) (Fig. 3b and c). But because abnormal movements are not observed in mutant mice with motor neuron degeneration^{16,17}, it is unlikely that loss of motor neurons could explain the phenotype of *TrkC* mutant mice. Whether damage to other structures such as the basal ganglia and joint and tendon afferents may contribute to this phenotype remains to be determined.

The limited lifespan of the mutant mice raises the possibility that these animals may have additional neuronal defects, although most central nervous system (CNS) structures known to express *trkC* transcripts^{2,5} appear to be grossly normal. Similar results have previously been obtained for mice lacking *TrkB*, the signalling receptor for brain-derived neurotrophic factor and NT4 (ref. 9). These observations contrast with the defects found in the peripheral nervous system of mice lacking *Trk* and *TrkB* receptors^{9,10}, and suggest that CNS neurons may be supported by more than one neurotrophin.

Finally, in humans, peripheral neuropathies which preferentially affect large diameter axons are sometimes associated with ataxia due to a positional sense loss and abnormal movements of the extremities known as pseudoathetosis^{18,19}. It is possible that these large-fibre neuropathies might be caused by a defect in NT3/*TrkC* signalling. □

Received 25 October 1993; accepted 25 February 1994.

1. Barbacid, M. in *Molecular Genetics of Nervous System Tumors* (eds Levine, A. J. & Schimdek H. H. 123–136 (Wiley, New York, 1993).
2. Lamballe, F., Klein, R. & Barbacid, M. *Cell* **66**, 967–979 (1991).
3. Ernfors, P., Merlio, J.-P. & Persson, H. *Eur. J. Neurosci.* **4**, 1140–1158 (1992).
4. Tessarollo, L. et al. *Development* **118**, 463–475 (1993).
5. Lamballe, F., Smeyne, R. & Barbacid, M. *J. Neurosci.* **14**, 14–28 (1994).
6. Lamballe, F., Tapley, P. & Barbacid, M. *EMBO J.* **12**, 3083–3094 (1993).
7. Tsoufas, R. et al. *Neuron* **10**, 975–990 (1993).
8. Valenzuela, D. M. et al. *Neuron* **10**, 963–974 (1993).
9. Klein, R. et al. *Cell* **75**, 113–122 (1993).
10. Smeyne, R. J. et al. *Nature* **368**, 246–249 (1994).
11. Willis, W. D. & Coggeshall, R. E. *Sensory Mechanisms in the Spinal Cord* (Plenum, New York, 1991).
12. Carroll, S. L. et al. *Neuron* **9**, 779–788 (1992).
13. Mu, X., Silos-Santiago, I., Carroll, S. L. & Snider, W. D. *J. Neurosci.* **13**, 4029–4041 (1993).
14. Snider, W. B., Zhang, L., Yusoof, X(7), Gorukanti, N. & Tsering, C. H. *J. Neurosci.* **12**, 3494–3508 (1992).
15. Kandel, E. R., Schwartz, J. H. & Jessell, T. M. *Principles of Neural Science* (Elsevier, New York, 1991).
16. Sendtner, M. et al. *Nature* **358**, 502–504 (1992).
17. Masu, Y. et al. *Nature* **365**, 27–32 (1993).
18. Griffen, J. W. et al. *Ann. Neurol.* **27**, 304–315 (1990).
19. Smith, B. E. *Neurol. Clinics* **19**, 735–758 (1992).
20. McBurney, M. W. et al. *Nucleic Acid. Res.* **20**, 5755–5761 (1991).
21. Mansour, S. L., Thomas, K. R. & Capocchi, M. R. *Nature* **336**, 348–352 (1988).
22. Joyner, A. L., Skarnes, W. C. & Rossant, J. *Nature* **338**, 153–156 (1989).
23. Chomczynski, P. & Sacchi, N. *Analyt. Biochem.* **162**, 156–159 (1987).
24. Marziani, G., Lazzaro, D. & Sorrentino, V. *Dev. Biol.* **157**, 182–190 (1993).

ACKNOWLEDGEMENTS. We thank A. Lewin, L. K. Long, M. Swerdel, J. D. Wallace and P. Zalamea for technical assistance, F. Lamballe for the mouse *trkC* cDNA clones, A. Reame for the mouse 129 genomic library, A. Schnapp for the pAS1 plasmid and G. Schnapp for the rRNA probe.

Co-evolution of ligand–receptor pairs

William R. Moyle, Robert K. Campbell*,
Rebecca V. Myers, Michael P. Bernard,
Yi Han & Xiaoying Wang

Department of Obstetrics and Gynecology,
University of Medicine and Dentistry of New Jersey—
Robert Wood Johnson (Rutgers) Medical School, Piscataway,
New Jersey 08854, USA

SPECIFIC receptors for lutropin^{1,2} (luteinizing hormone; LH) and follitropin^{3,4} (follicle-stimulating hormone; FSH) mediate the actions of human chorionic gonadotropin (hCG) and FSH⁵ on the gonads. Here we report that short independent sequences of the β -subunit enable hCG to distinguish between the receptors for FSH and LH. Residues between the 11th and 12th cysteines restrict FSH receptor binding; residues between the 10th and 11th cysteines and, to a much lesser extent, residues carboxy-terminal to the 12th cysteine also affect LH receptor binding. CF101–109, an hCG analogue containing hFSH β residues between the 11th and 12th cysteines, had high affinity for both LH and FSH receptors. Modifications to CF101–109 that reduce binding to either LH or FSH receptors yield gonadotropin analogues having differing ratios of LH:FSH activity. Ligand-binding specificity of the LH receptor is determined by residues encoded by parts of exons 2–4 and 7–9 which prevent hFSH binding but have little effect on hCG binding. FSH receptor specificity is controlled primarily by residues encoded by exons 5 and 6 that prevent hCG binding but have little effect on hFSH binding. These determinants can be interchanged to create receptor analogues that bind hCG and

* Present address: Ares Advanced Technology, 280 Pond Street, Randolph, Massachusetts 02368, USA.

TABLE 1 Receptor binding and activities of hormone and chimaera β -subunits as a function of amino-acid sequences between Cys(10) and the C terminus

Hormone or analogue	Sequence between Cys(10) and C terminus			Activity	
	Cys (10)	Cys (11)	Cys (12)	LH-RRA	FSH-RRA
hCG	—C ⁹³ RRSTTDC ¹⁰⁰	GGPKDHLTC ¹¹⁰	DDPR Q ¹⁴⁵ @	1.0	<0.008*
hFSH	—C ⁸⁷ DSDSTDC ⁹⁴	TVRGLGPSYC ¹⁰⁴	SFGEMKE@	<0.001*	1.0
CF94-117	—C ⁹³ DSDS	C ¹⁰⁰ TVRGLG	SYC ¹¹⁰ SFGEMKE@	<0.001*	0.53
CF94-114	—C ⁹³ DSDS	C ¹⁰⁰ TVRGLG	SYC ¹¹⁰ SFGE@	<0.02*	0.32
CF94-97, 108-114	—C ⁹³ DSDS	C ¹⁰⁰	SYC ¹¹⁰ SFGE@	0.15	<0.005*
CF94-97	—C ⁹³ DSDS	C ¹⁰⁰	 Q ¹⁴⁵ @	0.06	<0.006*
CF101-109	—C ⁹³	C ¹⁰⁰ TVRGLG	SYC ¹¹⁰ @	0.83	0.30
CF101-109Q ⁹⁴	—C ⁹³ Q	C ¹⁰⁰ TVRGLG	SYC ¹¹⁰ @	1.3	0.85
CF101-109Q ^{94,95}	—C ⁹³ QI	C ¹⁰⁰ TVRGLG	SYC ¹¹⁰ @	0.38	1.25
CF101-109D ⁹⁴ S ⁹⁵ D ⁹⁶	—C ⁹³ DSL	C ¹⁰⁰ TVRGLG	SYC ¹⁰⁰ @	0.038	0.83
CT101-117	—C ⁹³	C ⁹⁵ IHEAIKTYNC ¹⁰⁵	TKPKQSY@	0.33	<0.044*
eLH	—C ⁹³ QIK	C ¹⁰⁰ VFR Q	AC ¹¹⁰ APQA S ¹⁴⁹ @	0.53	0.18

Chimaeras were assembled using standard methods involving cassette mutagenesis and/or restriction fragment exchange as described for construction of CF94-114 and CF94-97 (ref. 17). The coding sequences of all constructs were confirmed by a double-stranded dideoxy DNA sequencing procedure using sequenase (USBC, Cleveland, OH) or with the Cycle Sequencing kit GIBCO-BRL (Bethesda, MD). The complementary DNA for these β -subunit analogues and the human α -subunit were expressed transiently in COS-7 cells downstream of the SV40 major late promoter as described¹⁷. Analogues secreted into serum-free DMEM medium were concentrated by ultrafiltration and their concentrations determined using monoclonal antibody sandwich immunoassays as described¹⁸, except that A113 (an antibody that binds the α -subunit of hCG) or B109 (an antibody that binds to a conformation unique to the β -subunit when it is combined with the α -subunit) was used to capture the analogues and ¹²⁵I-B105 (an antibody to the β -subunit) was used to detect them. Purified urinary hCG was used as a standard. LH receptor binding assays (termed LHR-RRA) were performed using ¹²⁵I-hCG and homogenates of rat corpora lutea and Chinese hamster ovary (CHO) cells expressing rat or human LH receptors as described¹⁷. FSHR-RRA were performed similarly using ¹²⁵I-hFSH and homogenates of bovine testes or CHO cells expressing human FSH receptors. The rat LH-receptor-expressing CHO cells were prepared by inserting the rat LH receptor cDNA obtained from a rat corpus luteum library¹⁹ downstream of a metallothionein promoter. The human LH-receptor-expressing CHO cells were prepared using a human LH receptor cDNA downstream of a cytomegalovirus promoter. Values shown are the averages of at least two assays and in most cases many more. They illustrate the abilities of the analogue to inhibit ¹²⁵I-hCG or ¹²⁵I-hFSH binding relative to that of hCG or hFSH. Except as noted, the values were obtained by dividing the concentration of hCG or hFSH that gave 50% inhibition of ¹²⁵I-hCG or ¹²⁵I-hFSH by the amount of analogue needed to give 50% inhibition. All analogues, except those shown with an asterisk, inhibited hCG and hFSH binding completely. Residues not shown are identical to those of the hCG β -subunit. '@' Refers to the location of the termination codon.

* No inhibition was observed by the highest amount of analogue available. Thus, the value indicated reflects the maximum possible activity; the actual activity is likely to be much lower.

hFSH. Our observations support a model in which distinct negative determinants restrict ligand-receptor interaction. This explains co-evolution of binding specificity in families of homologous ligands and their receptors. Natural or designed manipulation of these determinants leads to the 'evolution' of new, specific protein-protein interactions.

Apart from LH and hCG, which share a common receptor, human glycoprotein hormones bind their receptors with high specificity (0.001–0.1% crossreactivity)⁵. hCG/hFSH chimaeras containing hFSH β residues between cysteines 11 and 12 bound the FSH receptor with high affinity (Table 1) and elicited FSH signal transduction (Fig. 1). Sequences unique to hFSH β in this region are not required for FSH receptor binding, as equine LH (eLH) bound the FSH receptor well⁶ (Table 1). Analogues containing hCG β residues in this region failed to bind the FSH receptor, suggesting that residues between the 11th and 12th cysteines prevent hCG from binding to the FSH receptor. Chimaeras containing hFSH β residues between cysteines 10 and 11 had low LH activity, regardless of their sequences between cysteines 11 and 12 (Table 1). The influence of the hFSH β substitutions between cysteines 10 and 11 was due primarily to the replacement of positively charged residues normally found in lutropins, not a specific amino-acid sequence (Table 1; compare the activities of CF101-109, CF101-109Q⁹⁴, CF101-109Q^{94,95}, and eLH). Analogues with at least one positively charged amino acid in this region bound the LH receptor well. Those with the most negatively charged residues had the lowest affinities for the receptor. Maximal reduction in LH activity required negative charges between cysteines 10 and 11 and the presence of hFSH residues C-terminal of the 12th cysteine (Table 1, compare LH receptor binding of CF94-117, CF94-114 and CF101-109D⁹⁴S⁹⁵D⁹⁶).

The high-affinity ligand-binding sites of the LH and FSH receptors are located in their large extracellular domains^{7,8}. Triton X-100 extracts of COS-7 cells expressing an FSH receptor analogue containing LH receptor residues 1–201 in the extracellular domain (pMB68) bound hCG but not hFSH (Fig. 2a). Chimaeras containing more FSH receptor residues than pMB68 bound hFSH in proportion to the percentage of residues derived from the FSH receptor. As long as they contained LH receptor

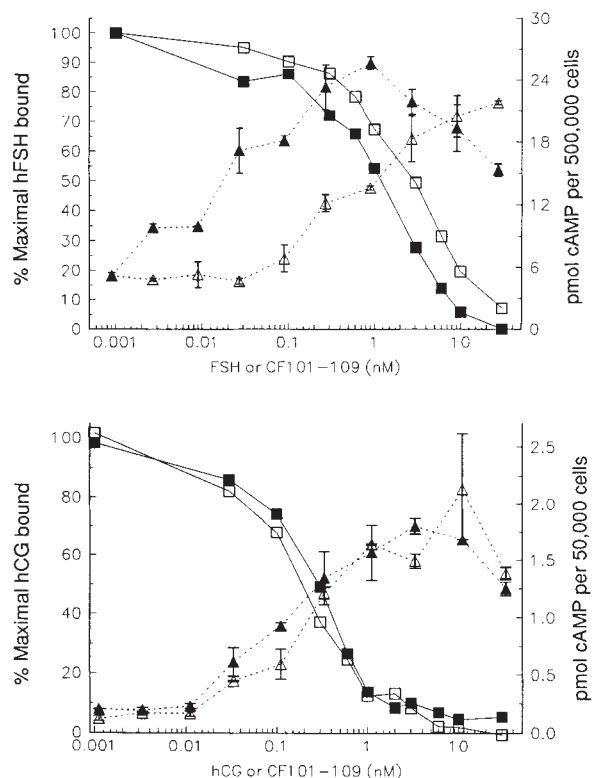


FIG. 1 Receptor binding and signal transduction activity of CF101-109. Inhibition of ¹²⁵I-hFSH binding and stimulation of cyclic AMP accumulation in cells expressing human FSH receptors (top) and inhibition of ¹²⁵I-hCG binding and stimulation of cyclic AMP accumulation in cells expressing human LH receptors (bottom). Procedures for measurement of receptor binding are described in Table 1 legend. CHO cells expressing the human FSH or LH receptors were incubated with the hormones or analogues for 15 min. The reactions were terminated and cyclic AMP measured by radioimmunoassay as described²⁰. Filled symbols, hCG or hFSH; open symbols, CF101-109; solid lines, receptor binding; broken lines, cyclic AMP accumulation.

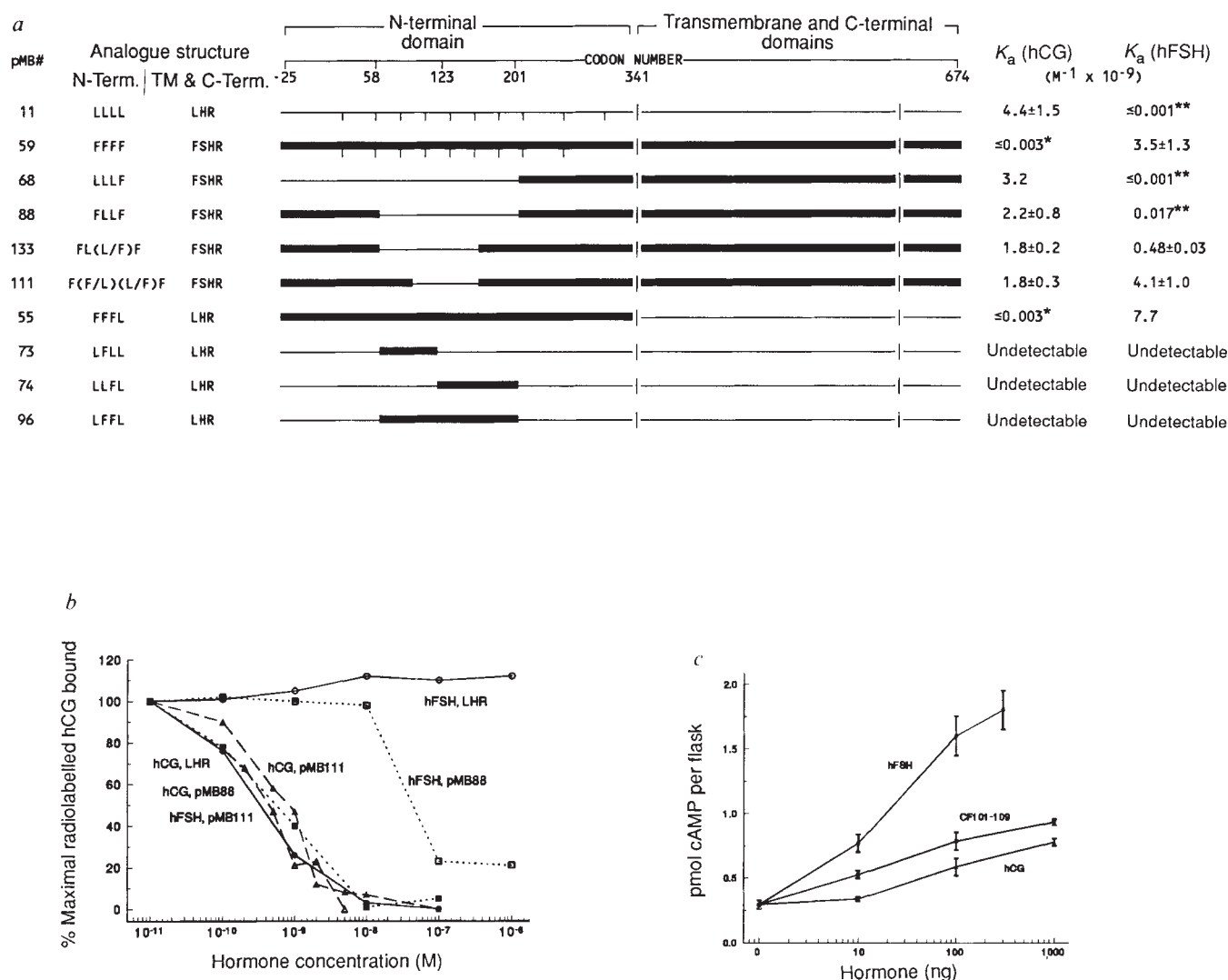


FIG. 2 **a**, Structure and activities of a few of the many receptor chimaeras used in this study. Chimaeras were expressed in COS-7 cells and their affinities for hCG or hFSH were calculated using the program LIGAND²¹. Previous studies have shown that some gonadotropin receptor analogues are retained within the cell^{8,22}. To minimize the influence of receptor expression on binding measurements, we solubilized the receptors and receptor chimaeras in Triton X-100 before measurement. K_a , Association constant. Values shown with a single or double asterisk were determined by measuring the ability of hCG or hFSH to inhibit binding of ^{125}I -hFSH or ^{125}I -hCG, respectively. Some receptor analogues bound neither ^{125}I -hCG nor ^{125}I -hFSH and their affinities for hCG and hFSH could not be determined ('Undetectable'). The chimaeras illustrated contained FSH or LH receptor (FSHR and LHR) transmembrane and cytoplasmic domains (TM & C-Term). The N-terminal (N-Term.) structures shown refer to the origins of receptor fragments used to build the chimaeras (L, LH receptor residues; F, FSH receptor residues). These are indicated by the numbers showing the approximate locations of endonuclease restriction sites. The thick bars and thin lines represent sequences derived from the FSH and LH receptor, respectively. The downward-facing ticks refer to the approximate junctions of the 10 exons of the FSH receptor and the 11 exons of the LH receptor. Note that only those corresponding to LH receptor residues 32, 56, 81, 106, 131, 157, 180, 205 and 267 are conserved in both receptors. Although they are highly homologous, the rat LH and FSH receptors have different numbers of residues. LH receptor residues 58, 93, 170 and 201 correspond to FSH receptor residues 60, 95, 171 and 202. **b**, The ability of hCG and hFSH to inhibit binding of ^{125}I -hCG to the LH receptor and selected chimaeras. **c**, Influence of hCG, hFSH and CF101-109 on signal transduction of pMB111.

METHODS. **a**, **b**, The sources of all radiolabelled and unlabelled hormones used in this study have been described previously^{17,23}. The

cDNA for the rat LH and FSH receptors was obtained by polymerase chain reaction (PCR) of a rat corpus luteum library prepared as described¹⁹. They were cloned into the *Xba*I-*Bam*HI sites in the poly-linker of pSVL (Pharmacia) and expressed in COS-7 cells⁸. During PCR isolation of rat FSH and LH receptor cDNA sequences from an ovarian corpus luteum library and in subsequent mutagenesis, we introduced restriction sites for *Xba*I (upstream of the initiation codon, -25), *Bgl*II (at codons 58 and 123), *Apa*I (at codon 202), *Mst*II (at codon 341) and *Bam*HI (downstream of the termination codon, 674). These codons refer to the rLH receptor sequence and the same sites were located at homologous positions in the rFSH receptor cDNA. Most chimaeras were prepared by exchanging restriction digestion fragments of the LH and FSH receptors. To make pMB111, we applied a multiple PCR procedure termed SOEing²⁴ using primers that had overlapping complementary sequences. The sequences of all constructs were confirmed by a double-stranded dideoxy method¹⁹. Procedures used to measure the binding of hormones to Triton X-100-solubilized receptors were identical to those reported⁸. **c**, Y-1 adrenal cells were transfected with a plasmid encoding the full-length pMB111 sequence downstream of a cytomegalovirus promoter and transfected cells were selected by their resistance to G418. The Y-1 cell line was chosen because it lacks the LH and FSH receptors and because it changes its shape (becomes round) in response to agents that induce cyclic AMP formation. Cell lines expressing pMB111 became round when treated with either hFSH or hCG (not shown), suggesting that pMB111 was able to transduce hCG and hFSH signals. This was confirmed by direct measurement of cyclic AMP as illustrated here. Note that the maximum response observed to hCG was always lower than that seen with hFSH. Similar cyclic AMP accumulation was also seen in COS-7 cells transiently expressing pMB111 (not shown).

residues 93–170, the chimaeras retained their high affinities for hCG, and one (pMB111) had high affinities for both hCG and hFSH. LH receptor analogues that contained FSH receptor sequences for part or all of the sequence at 93–170 did not bind either hormone (pMB73, pMB74 and pMB96). Thus, residues unique to the LH receptor in regions 1–93 and 170–201 do not seem to be needed for hCG binding, yet they prevent binding of hFSH to the LH receptor in a cooperative fashion. Conversely, residues unique to FSH receptor regions 95–171 are not needed for hFSH binding, but they prevent hCG from binding to the FSH receptor. This location of LH receptor specificity determinants is consistent with the abilities of hCG to bind truncated LH receptors⁷ and the abilities of hCG and thyroid stimulating hormone (TSH) to bind a TSH/LH receptor chimaera⁹. hCG and hFSH competed with each other for binding to pMB111 and other chimaeras (Fig. 2b), suggesting that residues common to both receptors make high-affinity contacts with the hormones. hFSH, hCG and CF101–109 stimulated cyclic AMP accumulation in COS-7 and Y-1 cells transfected with pMB111 (Fig. 2c). However, the efficacy of hCG was only 30–50% that of hFSH. Therefore, residues involved in ligand-binding specificity have little effect on signal transduction.

Comparisons of the LH receptor gene structure with the results of mutagenesis indicate that ligand-specific binding is encoded by exons 5 and 6, and possibly portions of exon 4. Similar comparisons for the FSH receptor indicate that exons 2–4 and 7–9 encode residues that enable it to exclude hCG. Thus, exon duplication and shuffling will explain the formation of a receptor that contains multiple specificity determinants. However, the coding sequences for specificity determinants of the gonadotropin β -subunit genes are located within a single exon, suggesting that they arose by a different mechanism.

Because of their essential roles in reproduction, the evolution of gonadotropins and their receptors must be highly coupled. Given the complexity of the hormones and their receptors, many mutations in either the hormones or the receptors would be expected to disrupt gonadal function and be lost quickly. Surprisingly, the specificities of some gonadotropins seemed to be reversed when their activities were compared in homologous and heterologous assays¹⁰. For example, as expected, FSH receptors from the sea turtle *Chelonia mydas* bind turtle FSH (Ch-FSH) better than turtle LH (Ch-LH). In contrast, snake (*Thamnophis*) FSH receptors bind Ch-LH and Ch-FSH equally well, and porcine FSH receptors recognize Ch-LH better than Ch-FSH¹⁰. Chicken LH binds to rat FSH receptors better than rat LH receptors, probably because of its sequence similarity to mammalian FSH in the region between cysteines 11 and 12 (ref. 11). Altered binding specificity also extends to interactions of mammalian hormones with mammalian receptors. Equine LH has high affinity for both rat LH and FSH receptors but in horses binds only to the LH receptor⁶. This suggests that binding specificity was flexible and varied independently of signal transduction activity during vertebrate evolution.

Structural similarities between the gonadotropin β -subunits and their genes¹² suggest that they evolved from a common precursor and then diverged functionally by acquiring residues that prevented the hormone heterodimers from binding to inappropriate receptors (Fig. 3). The receptors also seem to have arisen in this way^{2,4}. This pattern appears to be a general evolutionary mechanism that enables ligand–receptor couples to be created without requiring the simultaneous development of two new complementary binding sites. The evolution of lutropins, follitropins, and LH and FSH receptors seems to be an example in which specificity has reached a maximum. Many ligand–receptor pairs seem to be in different stages of this process, as indicated by the many examples of homologous ligands that bind a common receptor (for example hCG and hLH)⁵, homologous receptors that bind a common ligand (type A and B interleukin-8 receptors)¹³, or homologous ligands that share homologous receptors (for example, tumour necrosis factors α and β and

their p55 and p75 receptors)¹⁴. As illustrated by the reduced ability of human prolactin receptors to distinguish between prolactin and growth hormone compared with other prolactin receptors¹⁵, binding need not always evolve in the direction of higher specificity. Our observations demonstrate that specificity determinants can be manipulated independently in a rational fashion to create analogues with novel ligand-binding properties. Consequently, it should be possible to 'reverse' the steps of divergent evolution and 'restore' fully crossreactive binding to other homologous ligands, including those that, like hCG and hFSH, show very high receptor-binding specificity. In the case of the gonadotropins, this has permitted development of potent new analogues that stimulate ovarian development in rodents (G. J. Macdonald, Y.H., Yanhong Wang and W.R.M., unpublished data) and that may be useful for treating human infertility or for stimulating fertility in endangered species. Similarly, as found for tumour necrosis factor- α ¹⁶, it should be possible to accelerate divergent evolution or to reverse convergent evolution by incorporating inhibitory determinants into 'degenerate' ligands, thereby limiting their activities to a single receptor. □

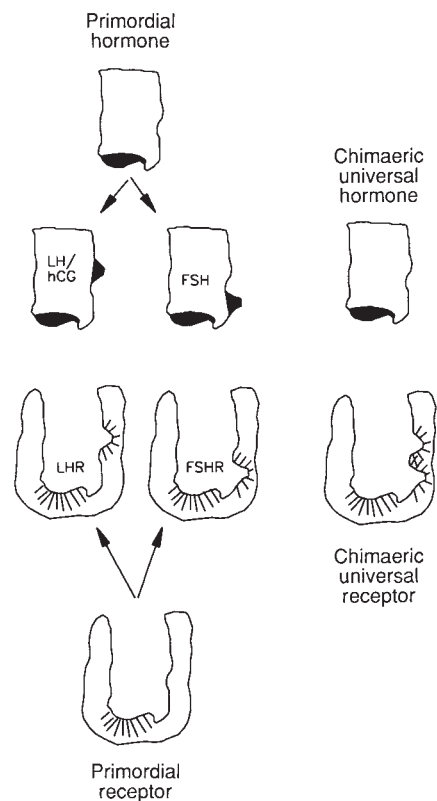


FIG. 3 Model of evolution of gonadotropin and gonadotropin receptor binding specificity. In this model a primordial gonadotropin β -subunit gene was duplicated. Receptor binding specificity was increased with the introduction of residues that blocked the binding of the hormone to inappropriate receptors. Replacement of these inhibitory residues yielded a glycoprotein hormone chimaera (CF101–109) that had lost its ability to distinguish between LH and FSH receptors. This model also postulates the existence of an ancestral receptor gene that was capable of binding multiple ligands. Following gene replication, specificity determinants were introduced into different regions of the receptors by mutations that restricted binding to one or more ligands. Thus, specificity was refined continuously and independently of affinity and/or ability to initiate signal transduction. These specificity determinants can be modified to create a 'universal' receptor. This model will also account for the evolutionary changes in the TSH receptor and many other families of homologous receptors and their ligands. However, *a priori*, there is no reason to assume that TSH specificity determinants will be located in the same region of the β -subunit as those that limit binding of hCG to the FSH receptor.

Received 22 September 1993; accepted 4 January 1994.

1. Segaloff, D. L., Sprengel, R., Nikolics, G. & Ascoli, M. *Rec. Prog. Horm. Res.* **46**, 261–301 (1990).
2. Koo, Y. B., Ji, I., Slaughter, R. G. & Ji, T. H. *Endocrinology* **128**, 2297–2308 (1991).
3. Sprengel, R., Braun, T., Nikolics, K., Segaloff, D. L. & Seeburg, P. H. *Molec. Endocr.* **4**, 525–530 (1990).
4. Heckert, L. L., Daley, I. J. & Griswold, M. D. *Molec. Endocr.* **6**, 70–80 (1992).
5. Pierce, J. G. & Parsons, T. F. A. *Rev. Biochem.* **50**, 465–495 (1981).
6. Murphy, B. D. & Martinuk, S. D. *Endocr. Rev.* **12**, 27–44 (1991).
7. Braun, T., Schofield, P. R. & Sprengel, R. *EMBO J.* **10**, 1885–1890 (1991).
8. Moyle, W. R., Bernard, M. P., Myers, R. V., Marko, O. M. & Strader, C. D. *J. Biol. Chem.* **266**, 10807–10812 (1991).
9. Nagayama, Y. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 902–905 (1991).
10. Licht, P. et al. *Rec. Prog. Horm. Res.* **33**, 169–248 (1977).
11. Noce, T. et al. *J. molec. Endocr.* **3**, 129–137 (1989).
12. Fiddes, J. C. & Talmadge, K. *Rec. Prog. Horm. Res.* **40**, 43–78 (1984).
13. Lee, J. et al. *J. Biol. Chem.* **267**, 16283–16287 (1992).
14. Vilcek, J. & Lee, R. H. *J. Biol. Chem.* **266**, 7313–7316 (1991).
15. Bewley, T. A. & Li, C. H. *Adv. Enzym.* **42**, 73 (1975).

16. VanOstade, X. et al. *Nature* **361**, 206–207 (1993).
17. Campbell, R. K., Dean Emig, D. M. & Moyle, W. R. *Proc. natn. Acad. Sci. U.S.A.* **88**, 760–764 (1991).
18. Moyle, W. R., Ehrlich, P. H. & Canfield, R. E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2245–2249 (1982).
19. Bernard, M. P., Myers, R. V. & Moyle, W. R. *Molec. cell. Endocr.* **71**, R19–R23 (1990).
20. Brooker, J., Harper, J. F., Terasaki, W. L. & Moylan, R. D. *Adv. Cyclic Nucleotide Res.* **10**, 1–33 (1992).
21. Munson, P. J. & Rodbard, D. *Analyt. Biochem.* **107**, 220–239 (1980).
22. Xie, Y.-B., Wang, H. & Segaloff, D. L. *J. Biol. Chem.* **265**, 21411–21414 (1990).
23. Moyle, W. R. et al. *J. Biol. Chem.* **265**, 8511–8518 (1990).
24. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. *Gene* **77**, 51–59 (1989).

ACKNOWLEDGEMENTS. We thank R. E. Canfield, and G. Bousfield for the purified urinary hCG and eLH used in these studies; S. Chappel for the purified recombinant hFSH and the CHO cell lines expressing the human FSH receptor, A. Hsueh for the gift of the human LH receptor vector, C. Kelly and M. Matzuk for helpful comments, and C. Kelly for showing us that the FSH receptor could be amplified from the corpus luteum cDNA library by PCR. These studies were supported by the NIH.

Recombinant G-protein $\beta\gamma$ -subunits activate the muscarinic-gated atrial potassium channel

Kevin D. Wickman*, Jorge A. Iñiguez-Lluhi†, Philip A. Davenport*, Ronald Taussig†, Grigory B. Krapivinsky*, Maurine E. Linder†, Alfred G. Gilman† & David E. Clapham*†

* Department of Pharmacology, Mayo Foundation, Rochester, Minnesota 55905, USA

† Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas 75235, USA

ACETYLCHOLINE activates inwardly rectifying potassium channels ($I_{K,ACH}$) in the heart¹ through muscarinic receptor binding and activation of pertussis-toxin-sensitive G proteins^{2,3}. Experiments showing that only the $\beta\gamma$ -subunit ($G\beta\gamma$) activates $I_{K,ACH}$ (ref. 4) were challenged by reports that only the activated α -subunit ($G\alpha$) was effective⁵. Here we examine $I_{K,ACH}$ regulation using purified brain and recombinant G-protein subunits. Six recombinant $G\beta\gamma$ -subunits activated $I_{K,ACH}$ with apparent half-maximal activation concentrations of 3–30 nM. Activation of $I_{K,ACH}$ by recombinant $G\alpha$ -GTP γ S was observed, but this was probably due to release of GTP γ S from the protein. Importantly, $I_{K,ACH}$ activity elicited by GTP γ S was inhibited by purified brain and recombinant $G\alpha$ -GDP, suggesting that native $G\beta\gamma$ plays a major role in this pathway. We conclude that $G\beta\gamma$ is a primary regulator of $I_{K,ACH}$ activity.

The first example of a $G\beta\gamma$ -activated effector was the muscarinic-gated atrial K^+ channel, $I_{K,ACH}$ (ref. 4). More recent results indicate that $G\beta\gamma$ is also important in other pathways (for a review, see ref. 6) As the list of roles for $G\beta\gamma$ grows, further knowledge of this complex is necessary to explain the specificity of signalling pathways.

Channels activated by purified $G\beta\gamma$ were identical to those stimulated by acetylcholine (pipette) and GTP (bath), or GTP γ S (bath) alone. Channel mean open times were the same in each case (~ 0.9 ms, not shown), and their current–voltage (I – V) relationships were superimposable (Fig. 1a and b). Experiments showing specific activation of $I_{K,ACH}$ by $G\beta\gamma$ are shown in Fig. 1c. Brain $G\beta\gamma$ generally activated $I_{K,ACH}$ within 30 s. The solubilizing detergent CHAPS did not activate $I_{K,ACH}$, but occasionally caused loss of patch integrity at higher concentrations (>100 μ M). Boiled $G\beta\gamma$ was inactive. To control for contamination

by free GTP γ S, we included 100 μ M GDP in the bath. GDP blocked or significantly delayed channel activation by up to 1 μ M free GTP γ S (not shown), but did not block or delay $G\beta\gamma$ -induced activity. Brain $G\alpha$ -GDP completely inhibited $I_{K,ACH}$ activity elicited by brain $G\beta\gamma$. As the channel was reactivated by high $G\beta\gamma$ concentrations, the inhibition was not due to patch vesicle formation or disruption of $I_{K,ACH}$ regulatory components. These results meet all reasonable criteria to demonstrate that $G\beta\gamma$ activates $I_{K,ACH}$.

We repeated this set of control experiments with one Sf9-derived $G\beta\gamma$, r $G\beta_{2\gamma 5}$. Isolation of highly pure, functional and homogeneous r $G\beta\gamma$ from Sf9 cells is possible through strong virus-driven foreign protein expression coupled with affinity purification of free r $G\beta\gamma$ s based on $G\alpha$ binding. Contamination by $G\alpha$ endogenous to Sf9 cells was not detected, and even if present, these subunits would be inactive (GDP-bound) because of the purification conditions^{7,8}. r $G\beta_{2\gamma 5}$ activated a channel identical to $I_{K,ACH}$ in mean open time and conductance. Activity in r $G\beta_{2\gamma 5}$ preparations was inactivated by boiling and unaffected by the absence or presence of GDP. Activity of $I_{K,ACH}$ elicited by r $G\beta_{2\gamma 5}$ was inhibited by *Escherichia coli*-derived myristoylated $G\alpha_{i3}$ -GDP (r $G\alpha_{i3}$ -GDP; Fig. 2a), which was shown to interact specifically with r $G\beta\gamma$ -subunits using several techniques^{7,8}. The apparent mean inhibitory concentration (K_i) for r $G\alpha_{i3}$ -GDP inhibition of channel activity was 10 nM (Fig. 2b).

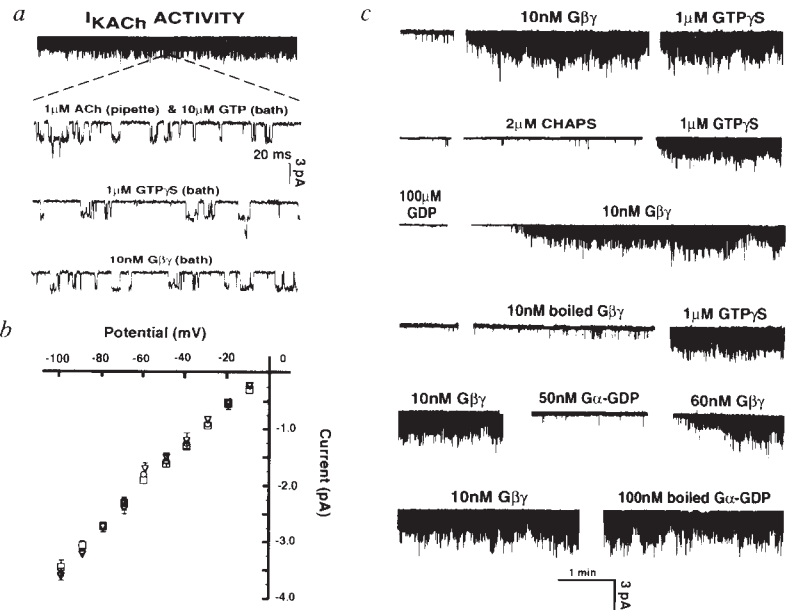
Figure 2c shows concentration–response curves for $I_{K,ACH}$ activation by various r $G\beta\gamma$ combinations. The six r $G\beta\gamma$ -subunit combinations showed few differences with respect to $I_{K,ACH}$ activation. r $G\beta_{1\gamma 1}$ was least potent, with an apparent half-maximal activation concentration (K_{act}) of 30 nM, while the other complexes had K_{act} s of 3–11 nM. The lower activity of r $G\beta_{1\gamma 1}$ compared with the other r $G\beta\gamma$ -subunits is due perhaps to differences in isoprenylation of $G\gamma$, a modification required for $G\beta\gamma$ membrane association⁹ and high-affinity $G\alpha$ (and effector) interaction⁷. Indeed, $G\gamma_1$ is predicted to be farnesylated, whereas $G\gamma_2$, $G\gamma_5$, and $G\gamma_7$ are predicted to be geranylgeranylated. Modification of $G\gamma$ in Sf9 cells is not, however, completely understood. Given the similarity in K_{act} s for this subset of r $G\beta\gamma$ -subunits we conclude that specificity is determined by other factors, perhaps specific interactions between $G\beta\gamma$ and certain $G\alpha$ subunits or receptor types.

GTP γ S applied to inside-out patches stimulates $I_{K,ACH}$ by activating (irreversibly) endogenous G proteins. This mode of channel stimulation was used to determine the identity of the native regulator of $I_{K,ACH}$. Results in Fig. 3 suggest that $G\beta\gamma$ is the primary mediator of $I_{K,ACH}$ activation in rat atrial myocytes. Brain $G\alpha$ -GDP (30–100 nM) rapidly inhibited channel activity elicited by 1 μ M GTP γ S, with channel reactivation on addition of 50–100 nM brain $G\beta\gamma$ (Fig. 3a). Similarly, the inhibition of $I_{K,ACH}$ activity by r $G\alpha_{i3}$ -GDP was concentration-dependent, with an apparent K_i of 3.0 nM (Fig. 3b). The inhibition seemed to be selective, as ongoing activity of another potassium channel,

† To whom correspondence should be addressed.

FIG. 1 Bovine brain $G\beta\gamma$ activates $I_{K,ACh}$ in inside-out patches from rat atrial myocytes. *a*, $G\beta\gamma$ (10 nM) activated channels identical to those stimulated by 1 μ M GTP γ S or acetylcholine (in the presence of GTP). *b*, I - V relationships for channels activated by 10 nM brain $G\beta\gamma$ (∇), 1 μ M GTP γ S ($\circ$) and 1 μ M acetylcholine (pipette) 10 μ M GTP (bath) ($\square$). *c*, $G\beta\gamma$ specifically activates $I_{K,ACh}$. In 28/30 experiments, 10 nM $G\beta\gamma$ activated $I_{K,ACh}$ within 3 min. Neither 200 μ M CHAPS ($n=30$) nor $G\beta\gamma$ samples boiled for 10 min ($n=8$) activated $I_{K,ACh}$. GDP (100 μ M) did not prevent activation by $G\beta\gamma$ ($n=9$). Bovine brain $G\alpha$ -GDP (30–100 nM) completely reversed activation of $I_{K,ACh}$ by 10 nM $G\beta\gamma$ ($n=8$), while boiled $G\alpha$ -GDP (100 nM) was ineffective ($n=9$). $G\beta\gamma$ (60–100 nM) reactivated channels inhibited by 30–100 nM $G\alpha$ -GDP ($n=8$).

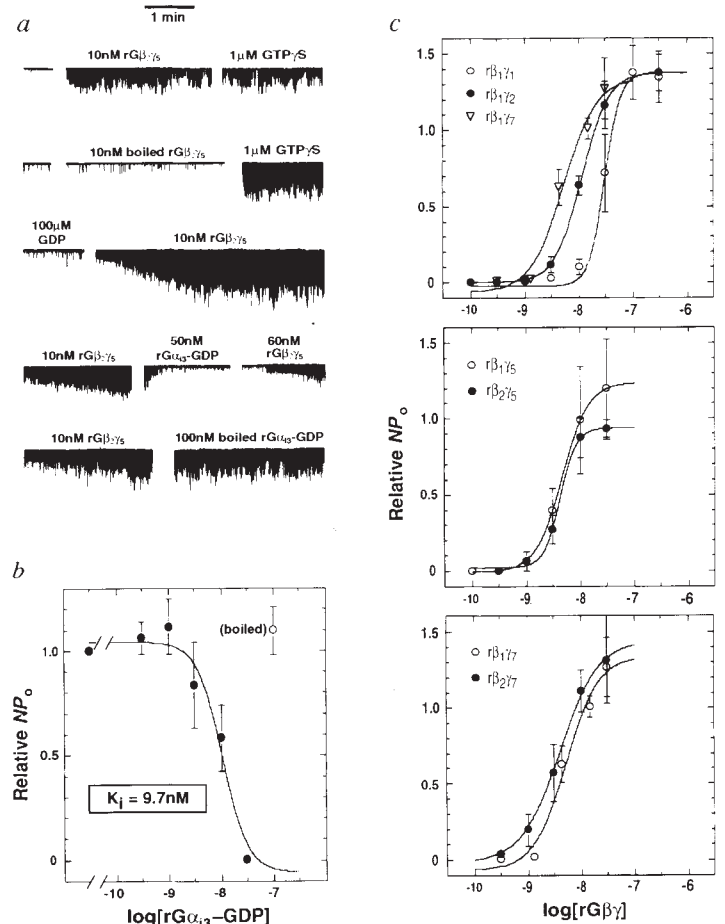
METHODS. Data were obtained from inside-out patches of isolated rat atrial cells⁴ using an EPC-7 patch-clamp amplifier. The pipette and bath solution was as follows (mM): K⁺ 133.5; Cl⁻ 133.5; EGTA 5; Mg²⁺ 2; HEPES 10; pH 7.2. Holding potential was -80 mV except for I - V determinations. Recordings were filtered at 2.5 kHz and displayed using a thermal recorder. Substances were prepared or diluted in recording solution and added manually to the bath. Spaces in data traces represent only the time necessary for substance addition. Patch viability was assessed after each experiment by application of GTP γ S. G-protein heterotrimers were isolated from bovine brain¹⁴, and subunits were separated as described¹⁵, followed by replacement of isolation detergents by CHAPS.



Subunits were not exposed to GTP γ S during any purification step. We verified that purified brain $G\alpha$ bound specifically to $G\beta\gamma$ immobilized on ω -aminobutyl agarose¹⁶. After diluting the subunits, CHAPS concentrations were 0.01–4 μ M.

FIG. 2 $rG\beta\gamma$ -subunits activate $I_{K,ACh}$. *a*, $rG\beta_2\gamma_5$ (10 nM) activated $I_{K,ACh}$ within 3 min in 26/26 attempts, while boiled $rG\beta_2\gamma_5$ was ineffective ($n=5$). GDP (100 μ M) did not prevent activation by 10 nM $rG\beta_2\gamma_5$ ($n=6$). $I_{K,ACh}$ activity elicited by 10 nM $rG\beta_2\gamma_5$ was completely inhibited by 30–70 nM $rG\alpha_{13}$ -GDP ($n=10$), with reactivation of the channel by 30–70 nM $rG\beta_2\gamma_5$ ($n=5$). Boiled $rG\alpha_{13}$ -GDP did not inhibit $rG\beta\gamma$ -induced channel activity ($n=10$). *b*, Concentration-dependence of $I_{K,ACh}$ inhibition by $rG\alpha_{13}$ -GDP ($n=6$). Activity at each concentration of $rG\alpha_{13}$ -GDP ($\bullet$) was normalized to activity induced by 10 nM $rG\beta_2\gamma_5$. Boiled 100 nM $rG\alpha_{13}$ -GDP ($\circ$) had no effect ($n=7$). *c*, Comparisons of the potencies of six $rG\beta\gamma$ -subunit combinations. K_{act} values for $rG\beta_1\gamma_1$, $rG\beta_1\gamma_2$, $rG\beta_1\gamma_5$, $rG\beta_1\gamma_7$, $rG\beta_2\gamma_5$, $rG\beta_2\gamma_7$, are (in nM): 30.0, 11.0, 4.6, 5.4, 4.1 and 3.9, respectively.

METHODS. Six $rG\beta\gamma$ combinations and *E. coli*-derived myristoylated $rG\alpha_{13}$ -GDP were prepared as described^{7,8,17} (CHAPS was exchanged for Lubrol after $rG\beta\gamma$ purification). In *b*, maximal activation by $rG\beta_2\gamma_5$ was determined by on-line monitoring, with each concentration of $rG\alpha_{13}$ -GDP given 2–2.5 min to exert an effect. Activity was quantified using the product NP_0 (N =number of channels; P_0 =open probability of 1 channel)⁴. NP_0 values taken during the last 1–1.5 min of each exposure were averaged and normalized to the activity of 10 nM $rG\beta_2\gamma_5$. In *c*, concentrations of $rG\beta\gamma$ were given 3 min to elicit maximal responses. $rG\beta\gamma$ -induced activity was normalized to 1 μ M GTP γ S. Occasionally, GTP γ S attenuated $rG\beta\gamma$ -induced channel activity, giving NP_0 values greater than 1.0. Protein concentration ranged over 100 pM–300 nM, giving CHAPS concentrations of 130 nM–130 μ M. NP_0 values were obtained using PCLAMP 5.1 software. Each point represents the average of 5–12 relative NP_0 values ($\pm$ s.e.m.).



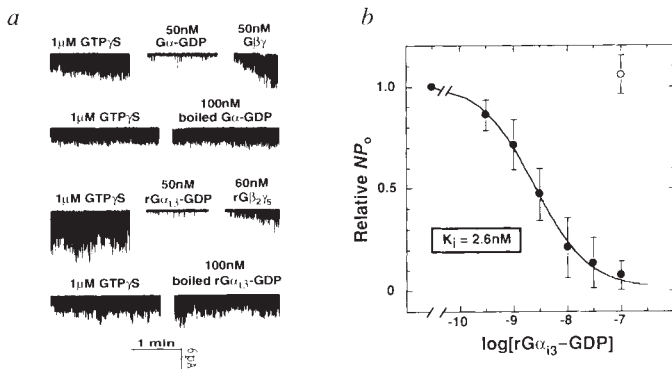
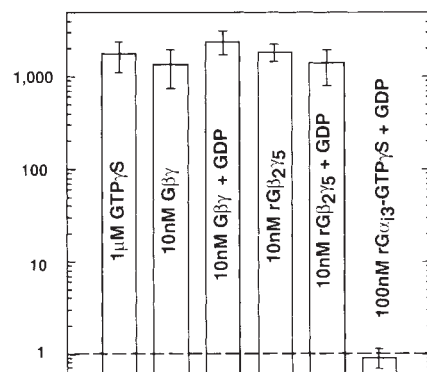


FIG. 3 $G\beta\gamma$ is the primary mediator of $I_{K,ACH}$ activation. *a*, Purified ($n = 11$) and recombinant ($n = 10$) $G\alpha$ -GDP subunits completely inhibit $I_{K,ACH}$ activity stimulated by 1 μ M GTP γ S. Reactivation of $I_{K,ACH}$ was observed with 50–70 nM $rG\beta_2\gamma_5$ in every experiment. Boiling for 8–10 min inactivated both preparations of $G\alpha$ -GDP subunits ($rG\alpha$: $n = 11$, $G\alpha$: $n = 9$). *b*, Concentration-dependence of reversal of $I_{K,ACH}$ activity by rG_{i13} -GDP ($n = 8$). NP_0 values for each concentration of rG_{i13} -GDP (●) were normalized to the activity of each patch after addition of GTP γ S. Boiled 100 nM rG_{i13} -GDP (○) had no effect.

FIG. 4 Comparison of the abilities of purified brain and recombinant G-protein subunits to increase $I_{K,ACH}$ activity within 3 min. GDP (100 μ M) was present in the bath where indicated. Values ($\pm$ s.e.m.) for the relative increase in $I_{K,ACH}$ activity were: 1 μ M GTP γ S, 1,770 ($\pm$ 660; $n = 11$); 10 nM $G\beta\gamma$, 1,810 ($\pm$ 600; $n = 9$); 10 nM $G\beta\gamma$ + GDP, 2,400 ($\pm$ 700; $n = 7$); 10 nM $rG\beta_2\gamma_5$, 2,100 ($\pm$ 410; $n = 11$); 10 nM $rG\beta_2\gamma_5$ + GDP, 1,400 ($\pm$ 600; $n = 5$); 100 nM rG_{i13} -GTP γ S + GDP, 0.907 ($\pm$ 0.2; $n = 9$).

METHODS. GTP γ S-bound (activated) $rG\alpha$ subunits were prepared as described¹⁴ and used either immediately after gel filtration (separation of activated protein from GTP γ S) or within 5 days of activation, with no intervening freeze-thaw cycle (storage on ice). $rG\alpha$ concentrations were determined by GTP γ S binding. NP_0 values for basal and stimulated (1.5–3 min after addition of substance to bath) conditions were determined as in Fig. 2b. Stimulated NP_0 values divided by basal NP_0 values were averaged to give fold increase ($\pm$ s.e.m.) in $I_{K,ACH}$ activity. The filtrate of a 100 nM sample of activated rG_{i13} (1 freeze-thaw cycle), obtained by centrifugation through a Millipore 10,000 NMWL filter unit, was tested in our assays and found to activate $I_{K,ACH}$ in 4/4 attempts. The amount of free and bound GTP γ S in samples passed through Bio-Rad 10 mL disposable desalting columns (size-dependent separation) was estimated. A sample of activated protein that had been frozen



(-70°C) and thawed once and one stored on ice for 72 h were examined. Radioactivity in the free GTP γ S peak was 5–20% of the total radioactivity in both samples. This corresponds to 5–20 nM free GTP γ S, a range of concentration that consistently activated $I_{K,ACH}$ in our hands.

$I_{K,ATP}$ was unaffected by $I_{K,ACH}$ inhibition (data not shown). Because rG_{i13} -GDP has been shown to interact with $rG\beta\gamma$, and brain $G\alpha$ -GDP specifically binds brain $G\beta\gamma$, the most parsimonious conclusion is that excess exogenous $G\alpha$ -GDP binds endogenous $G\beta\gamma$, forming inactive heterotrimers and inhibiting channel activity.

G_{i3} has been suggested to be the link between the receptor and $I_{K,ACH}$ (ref. 10). We found initially that 100 nM rG_{i13} -GTP γ S activated $I_{K,ACH}$, but not in the presence of 100 μ M GDP. Activation in the absence of GDP was reversed by rG_{i13} -GDP (100 nM, data not shown). These observations suggested that the rG_{i13} -GTP γ S preparation was contaminated by free GTP γ S. Indeed, the filtrate ($M_r < 10$ K) of a 100 nM sample of rG_{i13} -GTP γ S activated $I_{K,ACH}$. Size-exclusion chromatography confirmed that following one freeze-thaw cycle or extended incubation on ice, 5–20% of activated rG_{i13} released GTP γ S. rG_{i13} used immediately after activation and gel filtration (to separate bound and free GTP γ S) did not activate $I_{K,ACH}$ in the absence of GDP in 20 attempts (data not shown). This same preparation

was active, however, in adenylyl cyclase inhibition¹¹ (the assay is insensitive to nanomolar concentrations of GTP γ S), confirming protein viability in one effector system. Release of GTP γ S from activated $G\alpha$ cannot explain $I_{K,ACH}$ activation by picomolar $G\alpha$ concentrations reported previously⁵, as picomolar GTP γ S concentrations do not activate $I_{K,ACH}$ in the time frame of these experiments. Figure 4 summarizes the abilities of purified brain and recombinant G-protein subunits to activate $I_{K,ACH}$.

Using recombinant G-protein subunits, which have the advantage of purity and homogeneity, we show here consistent activation of $I_{K,ACH}$ by $G\beta\gamma$. Also, we have shown that both purified brain and recombinant $G\alpha$ -GDP inhibit $I_{K,ACH}$ activity elicited by $G\beta\gamma$, $rG\beta\gamma$ and GTP γ S. Inhibition of GTP γ S-induced activity by exogenous $G\alpha$ -GDP suggests a primary role for $G\beta\gamma$ in $I_{K,ACH}$ activation. Although these results contrast sharply with the lack of effect of rG_{i13} -GTP γ S, an interaction with $G\alpha$ and effector that modulates $I_{K,ACH}$ regulation is still conceivable. The recent cloning of GIRK1 (refs 12, 13), probably the channel examined here, will help our understanding of $I_{K,ACH}$ regulation. □

Received 28 October 1993; accepted 10 January 1994.

1. Sakmann, B., Noma, A. & Trautwein, W. *Nature* **303**, 250–253 (1983).
2. Pfaffinger, P. J., Martin, J. M., Hunter, D. D., Nathanson, N. M. & Hille, B. *Nature* **317**, 536–538 (1985).
3. Breitwieser, G. E. & Szabo, G. *Nature* **317**, 538–540 (1985).
4. Logothetis, D. E., Kurachi, Y., Galper, J., Neer, E. J. & Clapham, D. E. *Nature* **325**, 321–326 (1987).
5. Codina, J., Yatani, A., Grenet, D., Brown, A. M. & Birnbaumer, L. *Science* **236**, 442–445 (1987).
6. Clapham, D. E. & Neer, E. J. *Nature* **365**, 403–406 (1993).
7. Iñiguez-Lluhi, J., Simon, M. I., Robishaw, J. D. & Gilman, A. G. *J. Biol. Chem.* **267**, 23409–23417 (1992).
8. Ueda, N. *et al. J. Biol. Chem.* (in the press).
9. Muntz, K. H., Sternweis, P. C., Gilman, A. G. & Mumby, S. M. *Molec. Biol. Cell* **3**, 49–61 (1992).
10. Mattera, R. *et al. J. Biol. Chem.* **264**, 465–471 (1989).
11. Taussig, R., Quarmby, L. M. & Gilman, A. G. *Science* **268**, 9–12 (1993).
12. Kubo, Y., Reuveny, E., Slesinger, P. A., Jan, Y. N. & Jan, L. Y. *Nature* **364**, 802–806 (1993).
13. Dascal, N. *et al. Proc. natn. Acad. Sci. U.S.A.* **90**, 10235–10239 (1993).
14. Sternweis, P. C. & Robishaw, J. D. *J. Biol. Chem.* **259**, 13806–13813 (1984).
15. Roof, D. J., Applebury, M. L. & Sternweis, P. C. *J. Biol. Chem.* **260**, 16242–16249 (1985).
16. Pang, I.-H. & Sternweis, P. C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7814–7818 (1989).
17. Linder, M. E. *et al. J. Biol. Chem.* **266**, 4654–4659 (1991).

ACKNOWLEDGEMENTS. We thank S. Manahan for preparation of atrial myocytes, L. Krapivinsky for assistance in isolating the bovine brain G-protein subunits, and L. Stehno-Bittel and K. Hedin for reading the manuscript. This work was supported by NIH grants (A.G.G. and D.E.C.). D.E.C. is an Established Investigator of the American Heart Association and J.A.I.-L. is supported by a HHMI predoctoral fellowship.

Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer

C. Eric Bronner*, Sean M. Baker*, Paul T. Morrison†, Gwynedd Warren‡, Leslie G. Smith*, Mary Kay Lescoe§, Michael Kane||, Christine Earabino†, James Lipford||, Annika Lindblom¶, Pia Tannergård¶, Roni J. Bollag‡#, Alan R. Godwin‡#, David C. Ward‡**, Magnus Nordenskjöld¶, Richard Fishel§, Richard Kolodner||† & R. Michael Liskay*††

* Department of Molecular and Medical Genetics, L103, Oregon Health Sciences University, 3181 S.W. Sam Jackson Park Road, Portland, Oregon 97201-3098, USA

† Molecular Biology Core Facility, Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA

‡ Department of Genetics, and ** Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510, USA

§ Department of Microbiology and Molecular Genetics, Markey Center for Molecular Genetics, University of Vermont, Burlington, Vermont 05405, USA

|| Division of Cellular and Molecular Biology, Dana Farber Cancer Institute, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

¶ Department of Molecular Medicine, Karolinska Hospital, S-171 76 Stockholm, Sweden

THE human DNA mismatch repair gene homologue, *hMSH2*, on chromosome 2p is involved in hereditary non-polyposis colon cancer (HNPCC)^{1,2}. On the basis of linkage data, a second HNPCC locus was assigned to chromosome 3p21-23 (ref. 3). Here we report that a human gene encoding a protein, *hMLH1* (human MutL homologue), homologous to the bacterial DNA mismatch repair protein MutL, is located on human chromosome 3p21.3-23. We propose that *hMLH1* is the HNPCC gene located on 3p because of the similarity of the *hMLH1* gene product to the yeast DNA mismatch repair protein, *MLH1*^{4,5}, the coincident location of the *hMLH1* gene and the HNPCC locus on chromosome 3, and *hMLH1* missense mutations in affected individuals from a chromosome 3-linked HNPCC family.

An unidentified human gene located on chromosome 3p21-23 is involved in some HNPCC kindreds³. Examination of tumour DNA from the chromosome 3-linked kindreds revealed dinucleotide repeat instability similar to that observed for other HNPCC families⁶ and several types of sporadic tumours⁷⁻¹⁰. Because dinucleotide repeat instability is characteristic of a defect in DNA mismatch repair^{5,11,12}, HNPCC linked to chromosome 3p21-23 could result from a mutation in a second DNA mismatch repair gene. Repair of mismatched DNA in *Escherichia coli* requires a number of genes, including *mutS*, *mutL* and *mutH*. Defects in any one of these genes result in a general elevation of spontaneous mutation rates¹³. Studies with the yeast *Saccharomyces cerevisiae* have identified three DNA mismatch repair genes: a *mutS* homologue, *MSH2*¹⁴, and two *mutL* homologues, *PMS1*¹⁵⁻¹⁷ and *MLH1*⁴. Each of these three genes plays

an indispensable role in DNA replication fidelity, including the stabilization of dinucleotide repeats⁵. Because the DNA mismatch repair gene homologue *hMSH2* is the chromosome 2p HNPCC gene^{1,2,12}, other mismatch repair homologues, such as those related to *mutL*, are logical candidates for the HNPCC gene on chromosome 3.

To clone mammalian *MLH* genes, we used polymerase chain reaction (PCR) techniques such as those used to identify the yeast *MSH1*, *MSH2* and *MLH1* genes and the human *MSH2* gene^{1,2,4,14}. As template in the PCR, we used complementary DNA synthesized from poly(A)⁺-enriched RNA prepared from cultured primary human fibroblasts. The degenerate oligonucleotides we used were targeted at the amino-terminal amino-acid sequences KELVEN and GFRGEA (single-letter code; Fig. 1), two of the most conserved regions of the MutL family of proteins previously described for bacteria and yeast^{16,18,19}. Two PCR products of the expected size were identified, cloned, sequenced and shown to encode a predicted amino-acid sequence with homology to MutL-like proteins. These two fragments generated by PCR were used to isolate human cDNA and genomic DNA clones, one of which, *hMLH1*, is described here.

The *hMLH1* cDNA nucleotide sequence shown in Fig. 1 encodes an open reading frame 2,268 base pairs (bp). Figure 1 also shows the *hMLH1* protein, which consists of 756 amino acids and shares 41% identity with the protein product of the yeast DNA mismatch repair gene, *MLH1*⁴. The regions of the *hMLH1* protein most similar to yeast *MLH1* correspond to amino acids 11-317, showing 55% identity, and the last 13 amino acids, which are identical between the two proteins. Furthermore, the predicted amino-acid sequences of the human and mouse *MLH1* proteins show at least 74% identity (data not shown).

As a first step to determine whether *hMLH1* was a candidate for the HNPCC locus on human chromosome 3p21-23³, we used two separate *hMLH1* genomic fragments to map the *hMLH1* gene by fluorescence *in situ* hybridization (FISH)^{20,21}. Examination of several metaphase chromosome spreads localized *hMLH1* to chromosome 3p21.3-23 (Fig. 2). As independent confirmation of the location of *hMLH1* on chromosome 3, we used PCR with a pair of *hMLH1*-specific oligonucleotides and Southern blotting with a *hMLH1*-specific probe to analyse DNA from the NIGMS2 rodent/human cell panel (Coriell Institute for Medical Research, Camden, New Jersey, USA). Results of both techniques indicated chromosome 3 linkage (data not shown). We also mapped the mouse *MLH1* gene by FISH to chromosome 9 band E (data not shown); this is a position of synteny to human chromosome 3p²². Therefore, the *hMLH1* gene localizes to 3p21.3-23, within the genomic region implicated in chromosome 3-linked HNPCC families³.

Next we analysed blood samples from affected and unaffected individuals from two chromosome-3 candidate HNPCC families³ for mutations. Family 1 showed significant linkage (l.o.d. score=3.01 at recombination fraction of 0) between HNPCC and a marker on 3p. For family 2, the reported l.o.d. score (1.02) was below the commonly accepted level of significance, and thus only suggested linkage to the same marker on 3p. Linkage analysis of family 2 with the microsatellite marker D3S1298 on 3p21.3 gave a more significant l.o.d. score of 1.88 at a recombination fraction of 0 (unpublished data). Initially, we screened for mutations in two PCR-amplified exons of the *hMLH1* gene by direct DNA sequencing. In these two exons from three affected individuals of family 1 there were no differences from the expected sequence. In family 2, four individuals affected with colon cancer are heterozygous for a C to T substitution in an exon encoding amino acids 41-69, which corresponds to a highly conserved region of the protein (Fig. 3). For one affected individual, we screened PCR-amplified cDNA for additional sequence differences. The combined sequence information obtained from the two exons and cDNA of this one affected individual represents 95% (all but the first 116 bp) of the open

†† To whom correspondence should be addressed.

Present addresses: Institute for Molecular Medicine and Genetics, CB-2803 Sanders Research and Education Building Medical College of Georgia, Augusta, Georgia 30912, USA (R.J.B.); Department of Human Genetics, Bldg 533, Suite 5400, University of Utah, Salt Lake City, Utah 84112, USA (A.R.G.).

reading frame. We observed no nucleotide changes other than the C to T substitution. Also, four members of family 2, predicted by linkage data to be carriers (data not shown), and as yet unaffected, were heterozygous for the same C to T substitution. Two of these predicted carriers are below and two are

above the mean age of onset (50 yr) in this particular family. Two unaffected individuals examined from this same family, both predicted to be non-carriers (data not shown), showed the expected normal sequence at this position. Linkage analysis that utilizes the C to T substitution in family 2 gives a l.o.d. score of 2.23 at a recombination fraction 0. Using low-stringency cancer diagnostic criteria, we calculated a l.o.d. score of 2.53. These data indicate that the C to T substitution shows significant linkage to the HNPCC in family 2.

To determine whether this C to T substitution was a polymorphism, we sequenced the same exon amplified from the genomic DNA from 48 unrelated individuals and observed only the normal sequence. We have examined an additional 26 unrelated individuals using allele-specific oligonucleotide (ASO) hybridization analysis²³. None of these 74 unrelated individuals carry the C to T substitution. Therefore, the C to T substitution observed in family 2 individuals is not likely to be a polymorphism. We did not detect this same C to T substitution in affected individuals from the second chromosome 3-linked family, family 1³. We are continuing to study family 1 for mutations in *hMLH1*.

FIG. 1 Structure of human *MLH1*. a, The nucleotide sequence of the *hMLH1* cDNA and the amino-acid sequence of the protein predicted to be encoded by the *hMLH1* cDNA. The underlined DNA sequences are the regions of cDNA that correspond to the degenerate PCR primers that were originally used to amplify a portion of the *MLH1* gene (nucleotides 118–135 and 343–359). b, Alignment of the predicted human *MLH1* and *S. cerevisiae* *MLH1* protein sequences. Amino-acid identities are indicated by boxes, and gaps are indicated by dashes. c, Phylogenetic tree of MutL-related proteins. The phylogenetic tree was constructed using the predicted amino-acid sequences of seven MutL-related proteins: human *MLH1*; mouse *MLH1*; *S. cerevisiae* *MLH1*; *S. cerevisiae* *PMS1*; *E. coli* *MutL*; *S. typhimurium* *MutL* and *S. pneumoniae* *HexB*. The required sequences were obtained from GenBank release 7.3.

METHODS. Human *MutL*-related sequences were amplified by PCR using degenerate oligonucleotide primers (5'-CTTGATTCTAGAGC(T/C)-TCNCCNCT(G/G)(A/G)AANCC-3' and 5'-AGGTCCGAGCTCAA(A/G)GA(A/G)-(T/C)TNGTNGANAA-3'). The template for PCR was double-stranded cDNA synthesized from cultured primary human fibroblast poly(A)⁺ mRNA. PCR was in 50 µl containing cDNA template, 1.0 mM each primer, 5 IU of *Taq* polymerase (Cetus), 50 mM KCl, 10 mM Tris buffer pH 7.5 and 1.5 mM MgCl. PCR was for 35 cycles of 1 min 94 °C, 1 min at 43 °C and 1.5 min at 62 °C. Fragments of the expected size, ~212 bp, were cloned into pUC19 and sequenced. The cloned *MLH1* PCR product was labelled with a random primer labelling kit (RadPrime, Gibco BRL) and used to probe human cDNA and genomic cosmid libraries by standard procedures. DNA sequencing of double-stranded plasmid DNAs was as previously described¹. DNA sequence alignments and contiguous sequences were constructed using Sequencer 2.0.10 (Gene Codes Corporation, Ann Arbor, Michigan, USA). The pairwise protein sequence alignment was with DNASTar MegAlign (DNASTAR Inc., Madison, Wisconsin, USA) using the clustal method³⁰. Pairwise alignment parameters were ktuple of 1, gap penalty of 3, window of 5 and diagonals of 5. The phylogenetic tree was generated with the PILEUP program of the Genetics Computer Group software using a gap penalty of 3 and a length penalty of 0.1. The reported DNA sequence has been submitted to GenBank (accession number U07343).

a

```

CTTGGCTCTTCTGGCGCCAAATGTCGTTCTGGCAGGGGTTATTCTGGCGGCTGGACGAG 60
      M S F V A G V I R R L D E
ACAGTGGTGAACCGCATCGCGCGCGGGAAGTTATTCAGGCGGCGAGCTAATGCTATCAA 120
      T V V N R I A G E V I Q R P A N A T I
GAGATGATTGAGAACTGTTAGATGCTCAAAATCCACAAGTATTCAAGTGATTGTTAAAGAG 180
      E M I E N C L D A K S T S I Q V I V K E
GGAGGCCGGAAGTGATTGATGATCAAGACAAATGGCAGCGGATCGAGAAAGAGATCTG 240
      G G L K L I Q I Q D N G T G I R K E D L
GATATGTTATGTGAAGGTCTCACTACTAGTAACTGACGCTTGTGAGGATTTAGCCAGT 300
      D I V C E R F T T S T A K L Q S F E D L A S
ATTCTCACTATGGCTTTGAGGCTGAGGCTTTGGCCAGCATAGGCAATGTGGCTCATGTT 360
      I S T Y G F R G E A L A S I S H V A H V
ACTATTACAACAAAACAGCTGATGGAAGGTGTCATACAGAGCAAGTACTCAGATGGA 420
      T I T T K T A D G K C A Y R A S Y S D G
AAACTGAAAGCCCTCTTAAACCATGTGCTGGCAATCAAGGACCCAGATCAGCGTGGAG 480
      K L K A P P K P C A G N Q G T Q I T V E
GACCTTTTTCACATAGCCAGGAGGAGAAAGCTTTAAATAATCCAGTGAAGAAATAT 540
      D L F Y N I A T R K A L K N P S E E Y
GGGAAATTTTGGAGTTGTGGCAGGTATTCAGTACCAATGCGAGCATAGTTCTTCTCA 600
      G K I L E V Y G R Y S V H N A G I S F S
GTTAAAAACAAGGAGAGACAGTGTGATGTAGGACACTACCCAATGCTCAACCGTG 660
      V K K Q G E T V A D V R T L P N A S T V
GACAAATTCGCTCCATCTTGGAAATGCTGTAGTCGAGAAGTATAGAAATGATGATG 720
      D N I R S I F G N A V S R E L I E I G C
GAGGATAAAACCTAGCCTTCAAAATGATGGTACATATCAATGCAAACTCACTCACTG 780
      E D K T L A F K M N G Y I S N A N Y S V
AAGAATGATCTCTTCTACTCTCATCAACCATGCTGCTGGTGAAGTAACTTCTTGAGA 840
      K K C I F L L F I N H R L V E S T S L R
AAAGCCATAGAAACAGTGTATGAGCCTATTGGCCAAAACACACCCATCTCTGTAC 900
      K A I E T V Y A A Y L P K N T H P F L Y
CTCAGTTAGAAATCAGTCCCAAGATGGATGTTAATGTGACCCCAACAGCATGAA 960
      L S L E T S F Q N V D V M V H P T K H E
GTTCACCTCTGCGCAGGAGAGACCTCTGGAGCGGCTGACAGACATCGAGAGCAG 1020
      V H F L H E E S I L R V Q Q H I E S K
CTCCCTGGGCTCAATCTCTCAGGATGTACTTCAACAGACTTGTCTACAGGACTGCT 1080
      L L G S N S S R M Y F T Q T L L P G L A
GGCCCTCTGGGAGATGGTTAAATCCACAACAGTCTGACCTGCTTCTACTTCTGGA 1140
      G P S G E M V K S T S T L T S S T S G
AGTAGTGAAGGTCTATGCCACAGATGTTCTGATACAGATTCCCGGGAACAGAGCTT 1200
      S S D K V Y A H Q M V R T D S R E Q K L
GATGATTTCTGACGCTCTGAGCAAAACCCCTGCTCAGTCAAGCCAGGCGATTTGTCAC 1260
      A D F L Q P L S K P L S S Q P Q A I V T
GAGGATAGACAGATATTTCTAGTGGCAGGCTAGCAGCAGATGAGGAGATGCTTGA 1320
      E D K T D I S S G R Q Q D D E M L E
CTCCAGCCCTGCTGAAGTGGCTGCAAAATCAGAGCTTGGAGGGGATACACAAAG 1380
      L P A P A E V A A K N Q S L E G T T K
GGGACTTCAGAAATCTCAGAGAGAGAGAGGCTCTCCAGACCCGAGAGAGAGAT 1440
      G T S E M S E K R G P T S N E P R K H
CGGGAAGATTCTGATGTGAATGCTGGAAAGATGATCCGAAAGAAATGACTGCACT 1500
      R E D S D V E M V E D D S R K E M T A A
TGTAACCCCGGAGAGGATCATTAACCTCAGATGTTTGTAGTCTCCAGGAAGAAAT 1560
      C T P R R R I I N L T S V L S L Q E E I
AATGAGCAGGACATGAGGTTCTCCGGAGATGTTGCATACCACTCTCTGCTGGCTGT 1620
      N E Q G H E V L R E M L H N S F V G C
GTGAATCTCAGTGGGCTTGGCAGCAGCATCAACCAAGTTATACCTTCTCAACACC 1680
      V N P Q W A L A Q H Q T K L Y L L N T
AAGCTTAGTGAAGAACTGTTCTACAGCATCATTAATGATTGCTGAGGATTTGGTGT 1740
      K L S E E L F Y Q I L I Y D F A N F G V
CTCAGGTATCGGAGCCAGCCGCTTGTGACCTTGGCATGCTTGTCTTAGATGTCA 1800
      L R L S E P A F L F D L A M L A L D S P
GAGAGTGGCTGGACAGGAGATGTTGCCAAAGAGGACTTGTGTAATACATGTTGAG 1860
      E S G V E S G L E P L I D N Y V P L E G L
TTTCTGAGAAGAGGCTGAGATGCTGACAGATTTCTCTTGGAAATGATGAGGAA 1920
      F L K K K A E M L A D Y F S L E I D E
GGGAACCTGATGGATTACCCCTCTGATGACCAATATGTGCCCTTTGGAGGAGCTG 1980
      G N L I G L P L L I D N Y V P L E G L
CCTATTCTATCTTCCAGTACGACTGAGGTGATGGGACGAGAAAGAAAGAAATTTT 2040
      P I F I L R L A T E V N W D E E K E C F
GAAAGCCCTAGTAAAGATGCGCTATGTTCTATTCATCCGGAAGCAGTACATATCTG 2100
      E S L S K E C A M F Y I S I R K Q Y I S E
GAGTGACCTCTCAGGCGCAGCAGATGAGTGGCTGCTTCAATCCAACTCTCGGAG 2160
      E S T L S G Q Q S E V P S I P N S W K
TGGACTGTGGAACACATGTTCTATAAGCCCTTGGCTCACACATCTGCTCTTAAACAT 2220
      W T V E H I V Y K A L R S H I L P P K H
TTCACAGAGATGGAATATCTCGAGCTGCTTACCTGCTGATCTATACAAAGCTTT 2280
      F T E D C N I L Q L A N L P D L Y K V F
GAGAGGTGTAATATGTTTATTATGCACTGCTGGATGTTCTTCTTCTCTGCTATC 2340
      E R C
CGATACAAAGTGTGTTATCAAAAGTGTGATATCAAAAGTACCAACATAGTGTGGTAG 2400
      CACTTAAGACTATATCTGCTCTGATAGTATCTCTTTATACAGTGGATTGATTATA
AATAAATAGATGGCTTAAACATA 2460
      2484

```

b

```

HUMAN  M S F V A G V I R R L D E I Q R P A N A T I Q V I V K E D L G A T G T T A G A G C T A A T G C T A T C A A 100
YEAST  M S L R L I A L A S V A M I A G E V I Q R P A N A T I Q V I V K E D L G A T G T T A G A G C T A A T G C T A T C A A 97

HUMAN  G E A L A S I S V A H T T T P A D K C Y A K S T S I Q V I V K E D L G A T G T T A G A G C T A A T G C T A T C A A 200
YEAST  G E A L A S I S V A H T T T P A D K C Y A K S T S I Q V I V K E D L G A T G T T A G A G C T A A T G C T A T C A A 197

HUMAN  V A D V R T L P N A S T V G A G G T C T C A G A A T G C T G A T T C C G G G A A C A G A G C T T 296
YEAST  V A D V R T L P N A S T V G A G G T C T C A G A A T G C T G A T T C C G G G A A C A G A G C T T 296

HUMAN  H I S R O N V D N V H P T R E V F T H E S L E R V O H I E S K L G S N S R M Y F T C T L E G L A G P S G E M V K S T 380
YEAST  V I H R A V D N V H P T R E V F T H E S L E R V O H I E S K L G S N S R M Y F T C T L E G L A G P S G E M V K S T 396

HUMAN  H O M V T D S R E R L D A F T Q P I S K P L S S Q P Q A V T E D K T D I S S G R A R C D E E M L P A P A E V A A K N Q S L 479
YEAST  N K I L A L A S V A M I A G E V I Q R P A N A T I Q V I V K E D L G A T G T T A G A G C T A A T G C T A T C A A 481

HUMAN  E M V E D S R K E M T A A C T P R R R I I N L T S V L S L Q E E I A A T G A G C A G G A C A T G A G G T T C T C G G 571
YEAST  P S T A T G E A L P I S K D Y I R V K E S V M L T S K L G K V D S T R E L T D I F A L A V A M D E E R T A T H D K L P 581

HUMAN  V L R L S E P A P L P L A M L A L D S P L E G L T T C T A T T C T A T T C C A G T A C G A C T G A G G T G T T G A G 663
YEAST  G K I N Q S T N S V S I D V L Y L G S F P E L N D I A S K K L I S K I D M S S P N I M A S I V N D L L N D L K S V K S L 676

HUMAN  N I D E R E T F E S L S G A M F S I R K O Y I S E E S L G Q S V P G S I P N S K W T V E I V Y K A L R S H I L P P K H 756
YEAST  N I D E R E T F E S L S G A M F S I R K O Y I S E E S L G Q S V P G S I P N S K W T V E I V Y K A L R S H I L P P K H 769

```

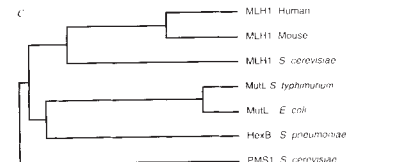


FIG. 2 Human *MLH1* localizes to 3p21.3-23 by fluorescence *in situ* hybridization. *a*, A metaphase spread showing hybridization of *hMLH1* is presented. Biotinylated *hMLH1* genomic probes were hybridized to banded human metaphase chromosomes as previously described^{20,21}. Detection was with fluorescein isothiocyanate (FITC)-conjugated avidin (green signal); chromosomes, shown in blue, were counterstained with 4',6-diamino-2-phenylindole (DAPI). Images were obtained with a cooled CCD camera, enhanced, pseudocoloured and merged with the following programs: CCD Image Capture; NIH Image 1.4; Adobe Photoshop and GeneJoin Maxpix, respectively. *b*, Composite of chromosome 3 from multiple metaphase spreads aligned with a human chromosome 3 ideogram. Region of hybridization (distal portion of 3p21.3-3p23) is indicated in the ideogram by a vertical bar.

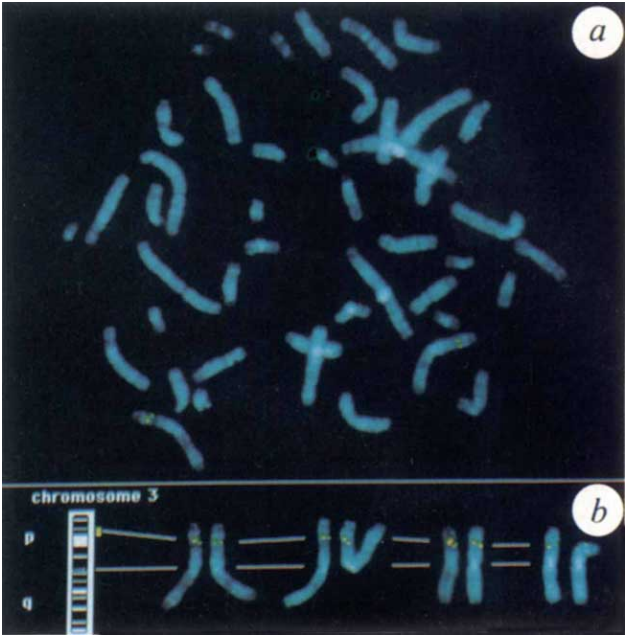
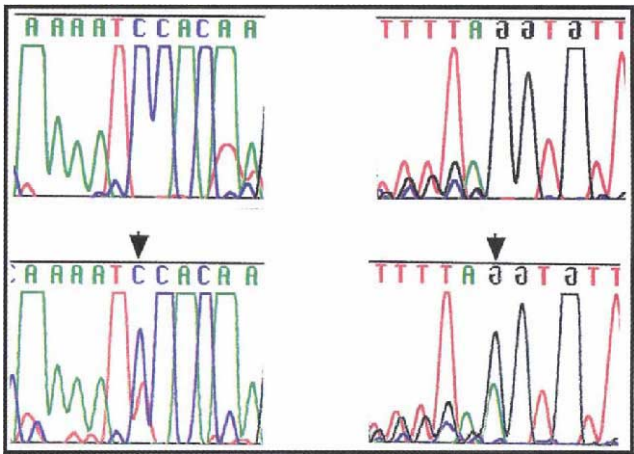


FIG. 3 Detection of *hMLH1* mutations in an HNPCC family showing linkage to chromosome 3p. *a*, Sequence chromatograms showing identification of a C to T substitution that produces a non-conservative amino-acid substitution at position 44 of the *hMLH1* protein. DNA samples obtained from the blood of individuals of interest were used as templates in PCR with primers (5'-GAAAGGTCCTGACTCTCC-3' and 5'-ATGTACATTAGAGTAGTTGC-3') that amplify the exon encoding amino acids 40-69. The PCR products were prepared, purified and analysed using an Applied Biosystems 377 sequencer as previously described¹. Sequence analysis of one unaffected (top, plus and minus strands) and one affected individual (lower, plus and minus strands) is presented. The position of the heterozygous nucleotide is indicated by an arrow. Analysis of the sequence chromatographs indicated that there is sufficient T signal in the C peak and enough A signal in the G peak for the affected individuals to be heterozygous at this site. Analysis of linkage of the C to T substitution in family 2 used standard methods³. Gene frequency is estimated at 0.001. Allele frequencies are assumed to be equal. All individuals with the mutation (all 4 affecteds, and 4 unaffecteds) share the same haplotype with markers on 3p21.3 (D3S643/D3S1478, D3S1298, D3S1029, Not73, D3S1100, D3S966, D3S1007). The markers have been mapped to this chromosomal region with human-chromosome-3-mouse hybrids and haplotype analysis in two chromosome 3-linked families (unpublished data). The mean age of onset of colon cancer in family 2 (50 yr) is higher than in family 1 and other HNPCC families^{25,26}. *b*, Amino-acid sequence alignments of the highly conserved region of the *MLH1* family of proteins surrounding the site of the predicted amino-acid substitution. Bold type indicates position of the predicted serine to phenylalanine amino-acid substitution in affected individuals, and the serine or alanine residues conserved at this position in MutL-like proteins. Dots indicate positions of highest amino-acid conservation. For the mouse *MLH1* protein the dots indicate that the sequence has not been obtained. Sequences were aligned as described for the phylogenetic tree in Fig. 1. For determination of the likely secondary structures for the normal primary sequence versus the



b

human <i>MLH1</i>	affected	VNRIAAGEVIQRPANAIKEMIENCLDAKFTSIQVIVKEGGLKLIQIQDNGTGIRKEDLDIVCER
human <i>MLH1</i>	normal	VNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLTQIQDNGTGIRKEDLDIVCER
mouse <i>MLH1</i>		PANAIKEMIENCLDAKSTNIQVVVKEGGLKLIQIQDNGTGIRKEDLDIVCER
<i>S. cerevisiae</i> <i>MLH1</i>		VNKIAAGEIIISPVNALKEMMENSIDANATMIDILVKEGGIKVLQITDNGSGINKADLPILCER
<i>S. cerevisiae</i> <i>PMS1</i>		VHRTISGQVITDLTTAVKELVDNSIDANANQIEIIFKDYGLEIECSNDNGGIDPSNYEFLALK
<i>E. coli</i> <i>MutL</i>		ANQIAAGEVVERPASVVKELVENSIDAGATRIDIDIERGGAKLIRIRDNGCGIKKDELALALAR
<i>S. typhimurium</i> <i>MutL</i>		ANQIAAGEVVERPASVVKELVENSIDAGATRVDDIERGGAKLIRIRDNGCGIKKDELALALAR
<i>S. pneumoniae</i> <i>HexB</i>		ANQIAAGEVIERPASVVKELVENALDAGSSQIIIEIEEAGLKKVQITDNGHGAHDEVELALRR
		

altered sequence in affected individuals from family 2, we used the Protean secondary structure module in the DNASTAR program (DNASTAR Inc.). The Chou-Fasman parameters were set at 103-alpha region threshold and 105-beta region threshold. Garnier-Robson calculated alpha and beta decision constants were both set to 0.

We suggest that the observed C to T substitution in the coding region of *hMLH1* represents the mutation that is the basis for HNPCC in family 2³. DNA sequence and ASO analyses did not detect the C to T substitution in 74 unrelated individuals. Thus the C to T substitution is not simply a polymorphism. The observed C to T substitution is expected to produce a serine to phenylalanine change at position 44 (Fig. 3). This amino-acid substitution is a non-conservative change in a conserved region of the protein (Figs 1 and 3). Secondary structure predictions using Chou–Fasman parameters (Fig. 3 legend) suggest a helix–turn–beta-sheet structure with position 44 located in the turn. The observed Ser to Phe substitution at position 44 lowers the prediction for this turn considerably, suggesting that the predicted amino-acid substitution alters the conformation of the *hMLH1* protein. Therefore, we propose that *hMLH1* represents a second DNA mismatch repair gene that is involved in HNPCC.

At present, we have no direct evidence that the *hMLH1* gene is involved in the correction of DNA mispairs. In bacteria and yeast, mutations affecting DNA mismatch repair cause increases in the rate of spontaneous mutation, including additions and deletions within dinucleotide repeats^{4,5,11,13,15,16,24}. In humans, mutation of *hMSH2* is the basis of chromosome-2 HNPCC^{1,2}, tumours of which show microsatellite instability and an apparent defect in mismatch repair¹². Chromosome 3-linked HNPCC is also associated with instability of dinucleotide repeats³. Combined with these observations, the high degree of conservation between the human *MLH1* protein and the yeast DNA mismatch repair protein *MLH1*, suggests that *hMLH1* is likely to function in DNA mismatch repair. During the isolation of the *hMLH1* gene, a second *MLH* gene, which does not map to chromosome 3, was identified and is predicted to encode a protein with strong similarity to yeast *PMS1* (our unpublished observations). This suggests that mammalian DNA mismatch repair, like that in yeast⁴, may require at least two MutL-like proteins.

We have described here a second DNA mismatch repair gene homologue, *hMLH1*, which is likely to be the hereditary non-polyposis colon cancer gene localized to chromosome 3p21–23³. Like other HNPCC families^{25,26}, chromosome 3p-families show apparent predisposition to several types of cancers³. The availability of the *hMLH1* and *hMSH2* gene sequences will aid the screening of HNPCC families for mutations in either gene. In addition, although loss of heterozygosity (LOH) of linked markers is not a feature of either the 2p or 3p forms of HNPCC^{3,6}, LOH involving the 3p21.3–23 region has been observed in several human cancers^{27–29}. Although speculative, this raises the possibility that *hMLH1* mutation may play some role in these tumours. Finally, defects in additional DNA mismatch repair genes may be the basis for the cancer in HNPCC families without chromosome 2p or 3p involvement³. □

21. Boyle, A., Feltquite, D. M., Dracopoli, N., Housman, D. & Ward, D. C. *Genomics* **12**, 106–115 (1992).
22. Lyon, M. F. & Kirby, M. C. *Mouse Genome* **91**, 40–80 (1993).
23. Wu, D. Y., Nozari, G., Schold, M., Conner, B. J. & Wallace, R. B. *DNA* **8**, 135–142 (1989).
24. Reenan, R. A. & Kolodner, R. D. *Genetics* **132**, 975–985 (1992).
25. Bishop, T. D. & Thomas, H. *Cancer Sur.* **9**, 585–604 (1990).
26. Lynch, H. T. et al. *Gastroenterology* **104**, 1535–1549 (1993).
27. Latif, F. et al. *Cancer Res.* **52**, 1451–1456 (1992).
28. Naylor, S. L., Johnson, B. E., Minna, J. D. & Sakaguchi, A. Y. *Nature* **329**, 451–454 (1987).
29. Ali, I. U., Lidereau, R. & Callahan, R. J. *nat. Cancer Inst.* **81**, 1815–1820 (1989).
30. Higgins, D., Bleasby, A. & Fuchs, R. *Comp. Apple Biosci.* **8**, 189–191 (1992).

ACKNOWLEDGEMENTS. We thank A. Harris, D.-M. Christie, M. Robatzek, K. Fish and T. Desai for technical assistance, J. Garber, F. Li and S. Verselis for control DNA samples, M. Litt and M. Forte for the NIGMS2 panel DNA, S. Friend for technical advice and T. Prolla for his comments on the manuscript. C.E.B. was a recipient of an American Cancer Society fellowship and S.M.B. received a fellowship from the C. G. Swabius Fund. This work was supported by NIH research grants to R.M.L., R.D.K., R.F. and D.C.W., a Cancer Center Core Grant to the D.F.C.I. from the NIH, and grants from Bert von Kantzow's Fund and the Stockholm Cancer Society. C.E.B. and S.M.B. contributed equally to this work.

Destabilization of the complete protein secondary structure on binding to the chaperone GroEL

Ralph Zahn*, Claus Spitzfaden†, Marcel Ottiger†, Kurt Wüthrich† & Andreas Plückthun*†

* Max-Planck-Institut für Biochemie, Protein Engineering Group, Am Klopferspitz, D-82152 Martinsried, Germany

† Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

PROTEIN folding *in vivo* is mediated by helper proteins, the molecular chaperones^{1–3}, of which Hsp60 and its *Escherichia coli* variant GroEL are some of the best characterized. GroEL is an oligomeric protein with 14 subunits each of *M*_r 60K^{4–6}, which possesses weak, co-operative ATPase activity^{7–9} and high plasticity¹⁰. GroEL seems to interact with non-native proteins, binding one or two molecules per 14-mer^{11–19} in a 'central cavity'²⁰, but little is known about the conformational state of the bound polypeptides. Here we use nuclear magnetic resonance techniques to show that the interaction of the small protein cyclophilin^{21,22} with GroEL is reversible by temperature changes, and all amide protons in GroEL-bound cyclophilin are exchanged with the solvent, although this exchange does not occur in free cyclophilin. The complete secondary structure of cyclophilin must be disrupted when bound to GroEL.

Exchange of cyclophilin amide protons (Fig. 1) in the presence of GroEL was studied under conditions in which cyclophilin could bind to GroEL, that is, at 30 °C and pH 6.0 (see below). An equimolar mixture of ¹⁵N-labelled cyclophilin and GroEL was heated in D₂O to 30 °C for 8 h to induce cyclophilin binding, followed by cooling to 6 °C for 14 h, to induce protein-chaperone dissociation (Fig. 2a). This cycle was repeated three times to ensure that most cyclophilin molecules were bound at least once even if the turnover of bound cyclophilin was slow (at pH 6.0 and 30 °C, only 50% of cyclophilin is bound to GroEL at equilibrium). Before the nuclear magnetic resonance (NMR) experiments, cyclophilin and GroEL were separated by cation-exchange chromatography in D₂O at 6 °C. The resulting cyclophilin fractions were pooled and concentrated to a protein concentration of 0.4 mM. A two-dimensional [¹⁵N, ¹H]-correlation ([¹⁵N, ¹H]-COSY) spectrum of this solution was completely empty (not shown), demonstrating that all amide protons had been exchanged with deuterium.

To show that the cyclophilin recovered in D₂O from the GroEL–cyclophilin complex is in the native folded form, it was

Received 8 February; accepted 28 February 1994.

1. Fishel, R. et al. *Cell* **75**, 1027–1038 (1993).
2. Leach, F. et al. *Cell* **75**, 1215–1225 (1993).
3. Lindblom, A., Tannergård, P., Werelius, B. & Nordenskjöld, M. *Nature Genet.* **5**, 279–282 (1993).
4. Prolla, T. A., Christie, D.-M. & Liskay, R. M. *Molec. cell. Biol.* **14**, 407–415 (1994).
5. Strand, M., Prolla, T. A., Liskay, R. M. & Petes, T. D. *Nature* **365**, 274–276 (1993).
6. Aaltonen, L. A. et al. *Science* **260**, 812–816 (1993).
7. Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. & Perucho, M. *Nature* **363**, 558–561 (1993).
8. Han, H.-J., Yanagisawa, A., Kato, Y., Park, J.-G. & Nakamura, Y. *Cancer* **53**, 5087–5089 (1993).
9. Risinger, J. I. et al. *Cancer* **53**, 5100–5103 (1993).
10. Thibodeau, S. N., Bren, G. & Shaid, D. *Science* **260**, 816–819 (1993).
11. Levinson, G. & Gutman, G. A. *Nucleic Acids Res.* **15**, 5323–5338 (1987).
12. Parsons, R. et al. *Cell* **75**, 1227–1236 (1993).
13. Modrich, P. A. *Rev. Genet.* **25**, 229–253 (1991).
14. Reenan, R. A. & Kolodner, R. D. *Genetics* **132**, 963–973 (1992).
15. Williamson, M. S., Game, J. C. & Fogel, S. *Genetics* **110**, 609–646 (1985).
16. Kramer, W., Kramer, B., Williamson, M. S. & Fogel, S. *J. Bact.* **171**, 5339–5346 (1989).
17. Bishop, D. K., Anderson, J. & Kolodner, R. D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3713–3717 (1989).
18. Mankovich, J. A., McIntyre, C. A. & Walker, G. C. *J. Bact.* **171**, 5325–5331 (1989).
19. Prudhomme, M., Martin, B., Mejean, V. & Claverys, J. J. *Bact.* **171**, 5332–5338 (1989).
20. Lichter, P. et al. *Science* **247**, 64–69 (1990).

† To whom correspondence should be addressed at: Biochem. Institut, Universität Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland.

re-exchanged into H₂O and kept at 26 °C for 2 weeks. A [¹⁵N, ¹H]-COSY spectrum of this sample (Fig. 2*d*) had the same appearance as a spectrum of native cyclophilin (Fig. 2*b*), except that the signals corresponding to the 39 most slowly exchanging amide protons (Fig. 1) were very weak. A [¹⁵N, ¹H]-COSY spectrum of cyclophilin treated as before (Fig. 2*a*) in the absence of GroEL showed that the 39 backbone amide protons which exchange most slowly in free cyclophilin²³ gave strong peaks. Residual peak intensities were observed for some residues with medium-fast exchange (Figs 1 and 2*c*), corresponding to the behaviour of free cyclophilin in D₂O solution without the special treatment used in Fig. 2*a*. Experiments with cross-sections²⁴ support these findings. The cross-section Fig. 2*b'* contains peaks from a rapidly and a slowly exchanging amide proton. In cross-section Fig. 2*c'* only the peak of the slowly exchanging amide proton is left. In cross-section Fig. 2*d'* the rapidly exchanging proton has been fully re-exchanged, whereas less than 10% of the slowly exchanging protons have been re-exchanged. A complete survey of the backbone amide proton exchange in free and GroEL-bound cyclophilin is given in Fig. 2*e, f*, showing that residues that were strongly protected against exchange in a solution of free cyclophilin in D₂O re-exchange only to a limited extent in the experiment shown in Fig. 2*d*.

To investigate the kinetic stability of the complex formed between cyclophilin and GroEL, a cyclophilin solution containing only 0.1 equivalents of GroEL (but otherwise identical to the solutions used in Fig. 2*c* and *d*) was subjected to the treatment of Fig. 2*a*. If the GroEL-cyclophilin complex is relatively stable under these conditions, one would predict a 14% reduction in the intensity of the cross peaks corresponding to slowly exchanging amide protons, since in each cycle of Fig. 2*a* only about half of GroEL would bind cyclophilin (Fig. 3*d*), leading to exchange. A much bigger intensity loss could result if cyclophilin exchanged in and out of the GroEL binding site. The [¹⁵N, ¹H]-COSY spectrum from this experiment (not shown) was very similar to that in Fig. 2*c*. Only a small exchange enhancement can thus be attributed to cyclophilin turnover, and the GroEL-cyclophilin complex at 30 °C and pH 6.0 must be kinetically rather stable.

When 0.005 stoichiometric equivalents of GroEL were present during the NMR observation of cyclophilin (much more would not be soluble at these high molar concentrations), the exchange of the slowly exchanging protons was not measurably affected (see above). There was significant line broadening in the NMR spectra however, which persisted after eightfold dilution of the protein sample, but disappeared on addition of 6 mM Mg-ATP

TABLE 1 Kinetics of cyclophilin folding

	pH	T-shift (°C)	[Cyclophilin] (μM)	[GroEL] (μM)	Rate constant
Aggregation	7.0	25→54	25	—	$1.3 \times 10^3 \text{ s}^{-1} \text{ M}^{-1}$
Unfolding	7.0	25→48	25	25	$2.3 \times 10^{-3} \text{ s}^{-1}$
	7.0	25→46	25	25	$1.9 \times 10^{-3} \text{ s}^{-1}$
	6.0	6→30	1	5	$7.6 \times 10^{-5} \text{ s}^{-1}$
Refolding	7.0	46→1	25	25	$2.9 \times 10^{-4} \text{ s}^{-1}$
	7.0	46→5	25	25	$5.5 \times 10^{-4} \text{ s}^{-1}$
	6.0	30→6	1	5	$1.8 \times 10^{-5} \text{ s}^{-1}$

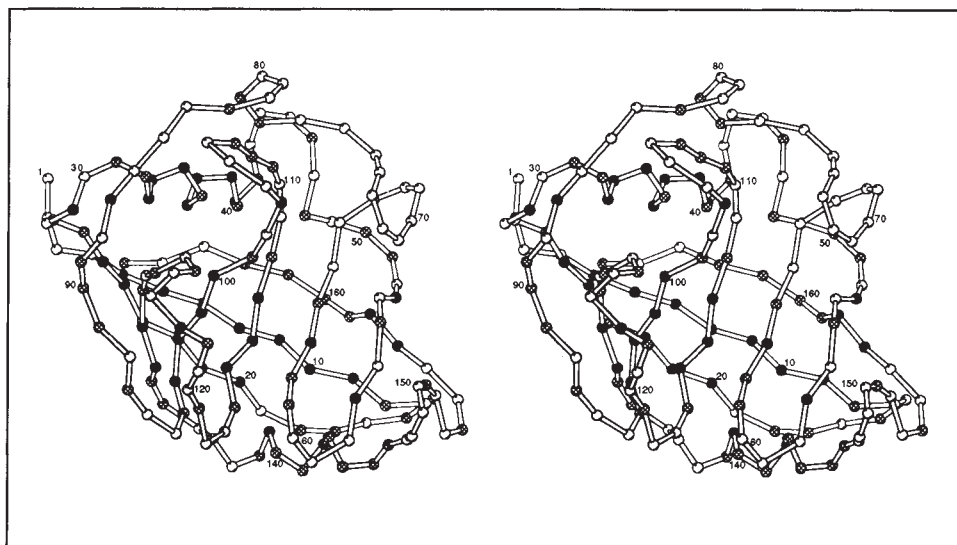
The folding kinetics were measured after shifting the temperature of a cyclophilin-containing solution to the value indicated and measuring the enzymatic activity as a function of time. The rate constants were determined by a three-parameter fit (starting activity, final activity, rate constant) to first-order kinetics (for GroEL-assisted unfolding and refolding) or second-order kinetics (for irreversible aggregation in the absence of GroEL).

(Fig. 2*g*). Under the conditions of these experiments (pH 6.0, 26 °C), the GroEL present contains some stably bound cyclophilin but is probably not saturated. Cyclophilin must thus interact transiently with either the free GroEL or the GroEL-cyclophilin complex, or both. While they are bound, these cyclophilin molecules would be restricted to the slow rotational tumbling of the GroEL-cyclophilin complex, causing the observed line broadening. The broadening in turn indicates that the cyclophilin stays bound for more than about 1 ns. Mg-ATP suppresses this interaction, presumably by causing a change in GroEL conformation⁹. The fact that this type of binding induces no exchange of the most stable amide protons (Fig. 1) indicates that it does not involve unfolding or destabilization of native cyclophilin, setting it apart from that seen in Fig. 2*d*.

The conditions for these experiments were selected on the basis of biochemical measurements of the kinetics and equilibria of cyclophilin folding in the presence of GroEL. Gel chromatography at pH 7.0 (data not shown) showed that micromolar concentrations of cyclophilin do not give rise to a stable complex with equimolar amounts of GroEL at room temperature, but that a significant proportion of cyclophilin co-elutes with GroEL once the temperature of the column is raised to 58 °C. If the solution is subsequently cooled to 25 °C, the complex of GroEL and cyclophilin dissociates again.

Measurements of enzymatic activity (Fig. 3*a*) show that there is concentration-dependent, irreversible unfolding of cyclophilin in the absence of GroEL, the kinetics of which are consistent with a second-order reaction (Table 1) as is typical of aggregation processes. In the presence of equivalent concentrations of GroEL, most of the cyclophilin is protected from irreversible

FIG. 1 Stereo view of the polypeptide backbone in the three-dimensional NMR solution structure of cyclophilin²⁸ with identification of residues with slowly exchanging amide protons. Backbone amide nitrogen positions represented by circles are connected by virtual bonds; every tenth residue is numbered. The black circles identify the previously identified²³ locations of the 39 residues for which complete amide proton exchange could not be achieved in free cyclophilin without irreversible denaturation of the protein. Shaded and open circles identify residues with medium and fast amide proton exchange, respectively²³. The plot was made with the program MOLSCRIPT²⁹.



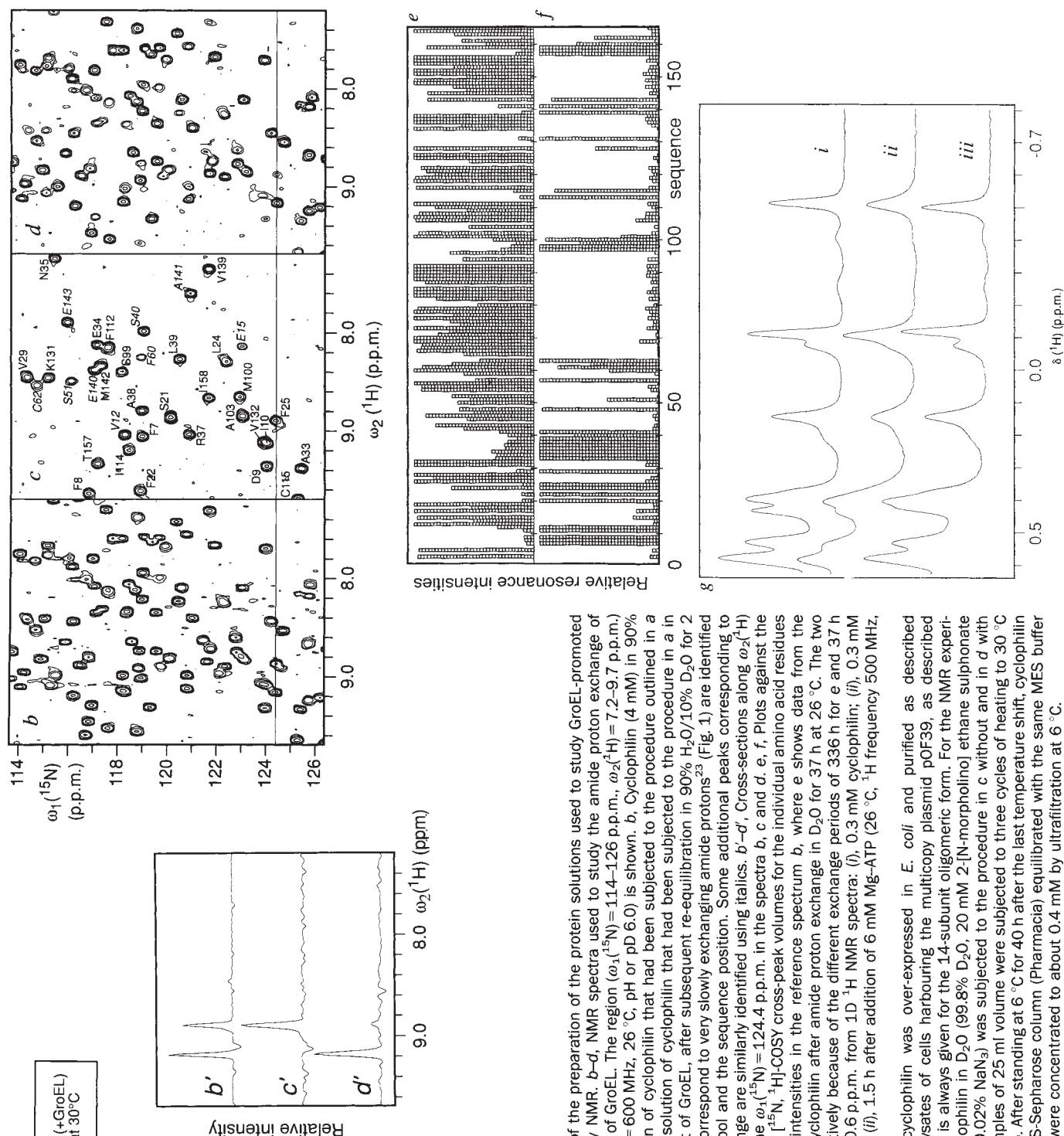
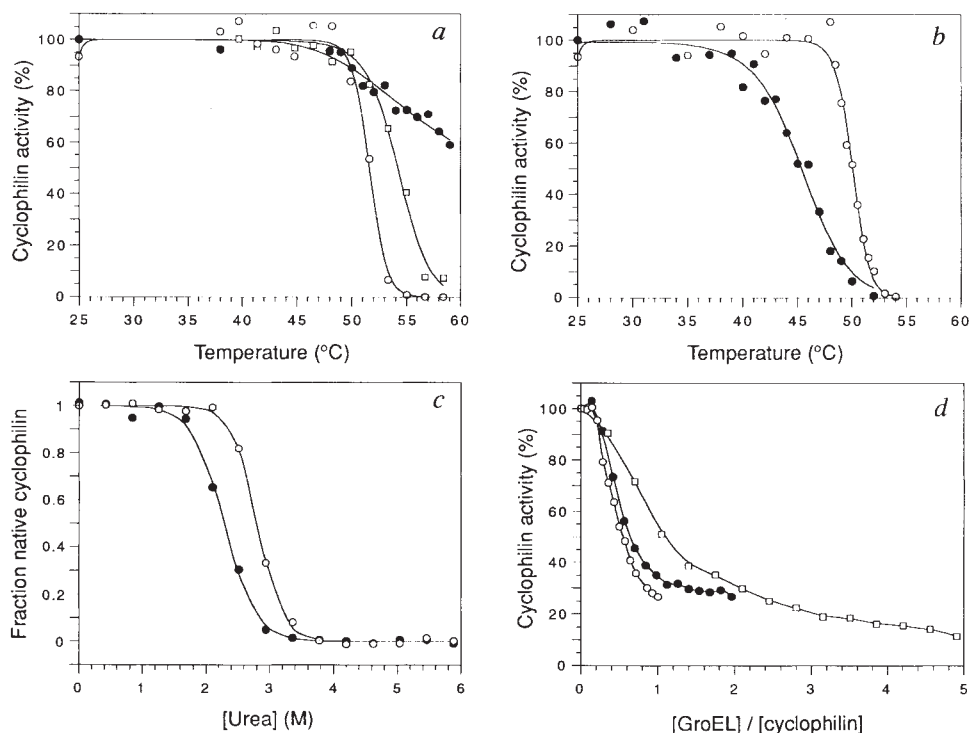


FIG. 2 **a**, Diagrammatic representation of the preparation of the protein solutions used to study GroEL-promoted amide proton exchange in cyclophilin by NMR. **b-d**, NMR spectra used to study the amide proton exchange of cyclophilin in the presence and absence of GroEL. The region ($\omega_1(^{15}\text{N}) = 114\text{--}126$ p.p.m., $\omega_2(^1\text{H}) = 7.2\text{--}9.7$ p.p.m.) of $[\text{I}^{15}\text{N}, ^1\text{H}]\text{-COSY}$ spectra (^1H frequency = 600 MHz, 26 °C, pH or pD 6.0) is shown. **b**, Cyclophilin (4 mM) in 90% $\text{H}_2\text{O}/10\%$ D_2O . **c**, A 0.5 mM D_2O solution of cyclophilin that had been subjected to the procedure outlined in **a** with no GroEL added. **d**, A 0.3 mM H_2O solution of cyclophilin that had been subjected to the procedure in **a** in the presence of a stoichiometric amount of GroEL, after subsequent re-equilibration in 90% $\text{H}_2\text{O}/10\%$ D_2O for 2 weeks at 26 °C. The cross peaks which correspond to very slowly exchanging amide protons²³ (Fig. 1) are identified in **c** with the amino acid one-letter symbol and the sequence position. Some additional peaks corresponding to amide protons with medium-slow exchange are similarly identified using italics. **b'-d'**, Cross-sections along $\omega_2(^1\text{H})$ taken at the position of the horizontal line, $\omega_1(^{15}\text{N}) = 124.4$ p.p.m. in the spectra **b**, **c** and **d**. **e**, **f**, Plots against the amino acid sequence of cyclophilin of the $[\text{I}^{15}\text{N}, ^1\text{H}]\text{-COSY}$ cross-peak volumes for the individual amino acid residues normalized by the corresponding peak intensities in the reference spectrum **b**, where **e** shows data from the experiment shown in **d** and **f** those for cyclophilin after amide proton exchange in D_2O for 37 h at 26 °C. The two data sets can be compared only qualitatively because of the different exchange periods of 336 h for **e** and 37 h for **f**. **g**, High-field region from -0.8 to 0.6 p.p.m. from 1D ^1H NMR spectra: **(i)**, 0.3 mM cyclophilin; **(ii)**, 0.3 mM cyclophilin + 1.5 μM GroEL; **(iii)**, same as **(ii)**, 1.5 h after addition of 6 mM Mg-ATP (26 °C, ^1H frequency 500 MHz, solvent 90% $\text{H}_2\text{O}/10\%$ D_2O).

METHODS. Recombinant ^{15}N -labelled cyclophilin was over-expressed in *E. coli* and purified as described previously²⁰. GroEL was purified from lysates of cells harbouring the multicopy plasmid pOF39, as described previously¹⁹. The concentration of GroEL is always given for the 14-subunit oligomeric form. For the NMR experiments **c** and **d**, a 20 μM solution of cyclophilin in D_2O (99.8% D_2O , 20 mM 2-[N-morpholino] ethane sulphonate (MES) pD 6.0, 2 mM DTT, 2 mM EDTA, 0.02% NaN_3) was subjected to the procedure in **c** without and in **d** with addition of 20 μM of GroEL. The two samples of 25 ml volume were subjected to three cycles of heating to 30 °C for 8 h and cooling to 6 °C for 14 h (see **a**). After standing at 6 °C for 40 h after the last temperature shift, cyclophilin was separated from GroEL on a HiLoad S-Sepharose column (Pharmacia) equilibrated with the same MES buffer in D_2O at 6 °C. The cyclophilin fractions were concentrated to about 0.4 mM by ultrafiltration at 6 °C.

FIG. 3 Folding of cyclophilin in the absence and presence of GroEL. *a*, Irreversibly unfolded cyclophilin after heating at pH 7.0. $\square$, low cyclophilin concentration (0.9 μ M) in the absence of GroEL; $\bullet$ and $\circ$, high cyclophilin concentration (9 μ M) in the presence or absence of GroEL (10 μ M), respectively. *b*, Portion of total (reversibly and irreversibly) inactivated cyclophilin after heating at pH 7.0 in the presence ($\bullet$) or absence ($\circ$) of GroEL. *c*, Urea denaturation curves of cyclophilin at 30 °C and pH 6.0 ($\bullet$) or pH 7.0 ($\circ$). *d*, Titration of cyclophilin with GroEL at 46 °C and pH 7.0 ($\bullet$, $\circ$), and at 30 °C and pH 6.0 ($\square$). The concentration of cyclophilin was 1 μ M ($\bullet$, $\square$) or 10 μ M ($\circ$).

METHODS. *a*, Cyclophilin was incubated for 20 min at the temperatures indicated in a buffer containing 100 mM K_2HPO_4/KH_2PO_4 pH 7.0 and 10 mM DTT. After cooling the samples to 10 °C and a subsequent waiting period of 1 h at 25 °C (to allow unfolded enzyme to refold), the enzymatic peptidyl-prolyl *cis-trans*-isomerase activity of cyclophilin was measured at 10 °C as described³¹. This experiment determines the sum of residual folded and refoldable cyclophilin. *b*, Solutions of 25 μ M cyclophilin in 100 mM K_2HPO_4/KH_2PO_4 at pH 7.0 and 10 mM DTT were incubated at various temperatures in the presence of 25 μ M GroEL or in the absence of GroEL. After 20 min the solutions were rapidly cooled to 1 °C by 100-fold dilution in the same buffer (to freeze the folding state of cyclophilin). The cyclophilin activity was immediately measured at 4.5 °C. In this experiment, only the amount of residual enzymatically active cyclophilin present in the incubation mixture is measured as a function of the incubation temperature. The remainder is reversibly or irreversibly inactivated. *c*, The enzymatic activity of a 1 μ M cyclophilin solution, incubated for 22 h at 30 °C at different urea concentrations, was measured at 4.5 °C. The solution was either at pH 6.0 (100 mM

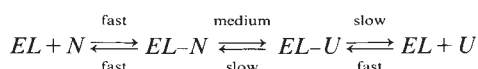


MES, 10 mM DTT) or at pH 7.0 (100 mM K_2HPO_4/KH_2PO_4 , 10 mM DTT). From these data, the fraction of native cyclophilin was calculated³² (note that these are not true equilibrium curves, however, because cyclophilin cannot be refolded after denaturation in 8 M urea under these conditions). *d*, Solutions of cyclophilin containing different concentrations of GroEL were incubated for 20 min at 46 °C in a buffer containing 100 mM K_2HPO_4/KH_2PO_4 , 10 mM DTT, pH 7.0, or for 19 h at 30 °C in a buffer containing 90 mM MES/38 mM K_2HPO_4/KH_2PO_4 , 10 mM DTT, pH 6.0. Immediately after heating, the enzymatic activity of cyclophilin, which is given relative to that of a GroEL-free cyclophilin solution, was measured at 4.5 °C.

denaturation (Fig. 3*a*), probably by forming a stable reversible complex with GroEL, as shown by the gel filtration experiments. In the absence of GroEL, the total unfolded portion of cyclophilin after heating at pH 7.0 (Fig. 3*b*) is similar to that in Fig. 3*a*, indicating that once cyclophilin is unfolded by heat, it cannot be folded back under these conditions. In the presence of GroEL, however, a reduction of cyclophilin activity occurs far below the denaturation temperature for free cyclophilin at a rate consistent with first-order kinetics (Table 1.) When the temperature was subsequently shifted back from 46 °C to 1 °C, cyclophilin recovered enzymatic activity with slow first-order kinetics (Table 1).

Urea denaturation curves show that cyclophilin is less stable at pH 6.0 than at pH 7.0 (Fig. 3*c*). The titration curves of cyclophilin with GroEL at 30 °C show that the net interaction between the chaperone and the substrate is stronger at pH 6.0 than at pH 7.0 (Fig. 3*d*), probably because of the lower thermodynamic stability of cyclophilin at pH 6.0, where it can be unfolded at 30 °C (Table 1). The close similarity of the titration curves obtained at pH 7.0 with different absolute concentrations of GroEL and cyclophilin indicates that most of the cyclophilin molecules may already be transiently bound to GroEL at the concentrations used; the observed transition in enzymatic activity would thus occur in the complex and be independent of concentration. Even at the highest GroEL concentrations used, the residual cyclophilin activity exceeded 20%.

These data can be interpreted in terms of the model:



where EL is GroEL, N is folded, enzymatically active cyclophilin with 39 exchange-protected amide protons, U is unfolded, enzymatically inactive cyclophilin with no exchange-protected amide protons, EL-N is the complex of EL and N, and EL-U is the complex of EL and U. The rates shown apply to the unfolding reaction observed at high temperature.

The exchange experiment of Fig. 2*d*, which started with native cyclophilin and led to tightly GroEL-bound cyclophilin, would thus start on the left, go to EL-U, and return to the left after the temperature had been lowered. In these experiments, cyclophilin was destabilized by lowering the pH, allowing unfolding to occur even at 30 °C (Table 1). At pH 7.0, cyclophilin remains stable at this temperature, but can be reversibly unfolded by GroEL at around 45 °C (Fig. 3*a, b*). By contrast, in the experiment with transiently bound cyclophilin (Fig. 2*g*) only the state EL-N is transiently occupied in the absence of ATP. The transient binding of native cyclophilin may be analogous to the previously reported stabilization of helical structure in a GroEL-bound peptide with a short residence time^{25,26}. *In vivo*, the reaction would presumably start on the right-hand side with U and EL, and as the $EL-N \rightleftharpoons EL-U$ equilibrium in the cell will normally be on the left, the reaction would proceed all the way from right to left, the last step being aided by ATP.

As cyclophilin is not bound instantaneously and quantitatively by GroEL under the conditions used (Fig. 3*d*), yet all protons are quantitatively lost, the protection factor of the amide protons labelled black in Fig. 1 must be reduced from $>10^7$ in native cyclophilin to $<10^3$ in GroEL-bound cyclophilin. By contrast, the protection factors of α -lactalbumin are 10^4 in the native state

and are lowered to 10^2 in the 'molten globule' state²⁷. Cyclophilin bound to GroEL thus seems to be fully unfolded, or in a low-energy molten globule state. The apparent absence of any stable secondary structure may be essential for the observed substrate promiscuity of GroEL¹¹⁻¹⁹. The chaperone may interact with interior side-chains to shift the equilibrium towards an unfolded state. By disrupting all native structure at least transiently, it may thus direct folding towards the native state by favouring stable structures and thereby avoiding aggregation, regardless of the final topology of the substrate. □

Received 22 September 1993; accepted 4 January 1994.

- Ellis, R. J. & van der Vies, S. M. *Rev. Biochem.* **60**, 321-347 (1991).
- Gething, M.-J. & Sambrook, J. *Nature* **355**, 33-45 (1992).
- Jaenicke, R. *Curr. Opin. Struct. Biol.* **3**, 104-112 (1993).
- Hendrix, R. W. *J. molec. Biol.* **129**, 375-392 (1979).
- Hohn, T., Hohn, B., Engel, A., Wurtz, M. & Smith, P. R. *J. molec. Biol.* **129**, 359-373 (1979).
- Hemmingsen, S. M. et al. *Nature* **333**, 330-334 (1988).
- Gray, T. E. & Fersht, A. R. *FEBS Lett.* **292**, 254-258 (1991).
- Bochkareva, E. S., Lissin, N. M., Flynn, G. C., Rothman, J. E. & Girshovich, A. S. *J. biol. Chem.* **267**, 6796-6800 (1992).
- Jackson, G. S. et al. *Biochemistry* **32**, 2554-2563 (1993).
- Zahn, R., Harris, J. R., Pfeifer, G., Plückthun, A. & Baumeister, W. *J. molec. Biol.* **229**, 579-584 (1993).
- Goloubinoff, P., Christeller, J. T., Gatenby, A. A. & Lorimer, G. H. *Nature* **342**, 884-889 (1989).
- Lamiet, A. A., Ziegehofer, T., Georgopoulos, C. & Plückthun, A. *EMBO J.* **9**, 2315-2319 (1990).
- Badcoe, G. et al. *Biochemistry* **30**, 9195-9200 (1991).
- Buchner, J. et al. *Biochemistry* **30**, 1586-1591 (1991).
- Höll-Neugebauer, B., Rudolph, R., Schmidt, M. & Buchner, J. *Biochemistry* **30**, 11609-11614 (1991).
- Martin, J. et al. *Nature* **352**, 36-42 (1991).
- Mendoza, J. A., Rogers, E., Lorimer, G. H. & Horowitz, P. M. *J. biol. Chem.* **266**, 13044-13049 (1991).
- Vitonen, P. V., Donaldson, G. K., Lorimer, G. H., Lubben, T. H. & Gatenby, A. A. *Biochemistry* **30**, 9716-9723 (1991).
- Zahn, R. & Plückthun, A. *Biochemistry* **31**, 3249-3255 (1992).
- Langer, T., Pfeifer, G., Martin, J., Baumeister, W. & Hartl, F.-U. *EMBO J.* **11**, 4757-4765 (1992).
- Handschumacher, R. E., Harding, M. W., Rice, J., Druggie, R. J. & Speicher, D. W. *Science* **226**, 544-547 (1984).
- Schreiber, S. L. *Science* **251**, 283-287 (1991).
- Wüthrich, K., Spitzfaden, C., Memmert, K., Widmer, H. & Wider, G. *FEBS Lett.* **285**, 237-247 (1991).
- Wüthrich, K. *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
- Landry, S. J. & Gierasch, L. M. *Biochemistry* **30**, 7359-7362 (1991).
- Landry, S. J., Jordan, R., McMacken, R. & Gierasch, L. M. *Nature* **355**, 455-457 (1992).
- Chyan, C.-L., Wormald, C., Dobson, C. M., Evans, P. A. & Baum, J. *Biochemistry* **32**, 5681-5691 (1993).
- Spitzfaden, C., Wider, G., Widmer, H. & Wüthrich, K. *Abstr. XV Int. Conf. Magnetic Resonance in Biological Systems*, Jerusalem, 192 (1992).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946-950 (1991).
- Weber, C. et al. *Biochemistry* **30**, 6563-6574 (1991).
- Fischer, G., Bang, H. & Mech, C. *Biomed. biochim. Acta* **43**, 1101-1111 (1984).
- Pace, C. N. *Trends Biotechnol.* **8**, 93-98 (1990).

ACKNOWLEDGEMENTS. We thank K. Memmert (Sandoz) for the cyclophilin expression plasmid and discussions about the production of ¹⁵N-labelled protein. The work was supported by the Deutsche Forschungsgemeinschaft, the Schweizerischer Nationalfonds and Sandoz, and a predoctoral fellowship from the Fonds der chemischen Industrie (to R.Z.).

A large-conductance mechanosensitive channel in *E. coli* encoded by *mscL* alone

Sergei I. Sukharev*, Paul Blount*, Boris Martinac*†, Frederick R. Blattner‡ & Ching Kung*§

* Laboratory of Molecular Biology and † Department of Genetics, University of Wisconsin-Madison, 1525 Linden Drive, Madison, Wisconsin 53706, USA

ALL cellular organisms respond to vibration, touch, gravity or changes in osmolarity, although the molecules on which such mechanosensations depend are unknown. Candidates include certain channels that gate in response to membrane stretch^{1,2}. Patch-clamp experiments with *Escherichia coli* envelope have revealed a mechanosensitive channel with very large conductance (MscL) and one with a smaller conductance (MscS)³⁻⁶ which may be important in osmoregulation. Here we have solubilized and fractionated the envelope, reconstituted the MscL activity *in vitro*, and traced it to a small protein, whose gene, *mscL*, we then cloned. Insertional disruption of *mscL* removes the channel activity, whereas re-expression of *mscL* borne on an expression plasmid restores it. MscL-channel activities were observed in material from a cell-free expression system with *mscL* as the only template. The *mscL* nucleotide sequence predicts a unique protein of only 136 amino acids, with a highly hydrophobic core and very different from porins or other known proteins.

A ramp of suction applied to a patch excised from an *E. coli* giant spheroplast first activates a 1-nS conductance from MscS channels with a sustained open state (Fig. 1a, asterisk) and then, at higher suction, a 3-nS flickering conductance of the MscL channels (Fig. 1a, inverted triangle). We have shown that the activities of these channels are intact after the membranes are

solubilized in octylglucoside and reconstituted into soybean-azolectin liposomes⁷. Our results show that these mechanosensitive channels are directly sensitive to stretch forces in the lipid bilayer^{5,6}. These prominent and distinct activities encouraged us to enrich and identify the channel proteins by following their activities *in vitro*.

Starting with total *E. coli* membranes, two different series of chromatographic enrichment each resulted in a relatively simple fraction with only one observed common component, a protein of relative molecular mass ~17,000 (*M_r* 17K) as judged by its mobility in SDS-PAGE (Fig. 1c, triangle). This protein was accompanied by another that migrated slightly faster in some fractions (not shown). When reconstituted into liposomes, these semipurified fractions had high MscL activity and were devoid of MscS activity (Fig. 1b). The 17K target protein was electroblotted onto an Immobilon membrane and microsequenced to reveal the sequence of 37 N-terminal amino-acid residues. A 3-kilobase *Pst*I genomic fragment at minute 72 of the chromosome had been found previously⁷ that encoded TrkA and another protein of unknown function. A database search matched our N-terminal MscL sequence to the partial sequence of this unknown protein encoded 3' of *trkA*. Therefore, using a λ clone in the Daniel and Blattner collection⁸, we subcloned this *Pst*I fragment and sequenced the entire *mscL*.

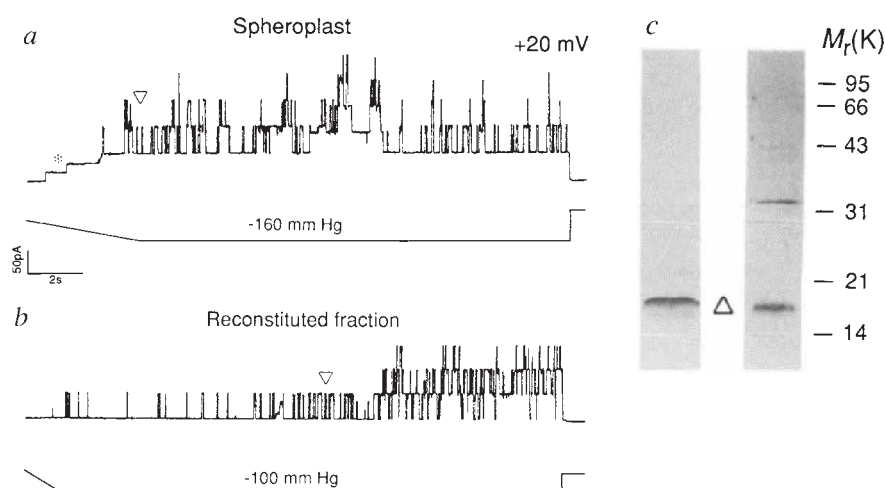
The open reading frame (ORF) of *mscL* predicts a protein of 136 amino-acid residues, with an *M_r* of 14,957, assuming no post-translational modifications (Fig. 2a). The complete sequence currently does not match any other protein or DNA sequence in several databases. A hydropathy analysis (Fig. 2b) shows that the C-terminal quarter of the peptide is hydrophilic, the remainder of the molecule being apparently very hydrophobic apart from two small stretches at residues 1-11 and 52-67. This hydrophobic core, covering most of the first three-quarters of the molecule, could encompass several transmembrane domains. However, MscL does not contain the multiple amphipathic β -domains that typically traverse the membrane in bacterial porins⁹, or in the mitochondrial voltage-dependent anion channel¹⁰ (R. Guy, personal communication). The elution profile of MscL activity from a calibrated gel-filtration column under non-denaturing conditions corresponded to *M_r* ~70K⁵, which may be indicative of multimerization.

† Present address: Department of Pharmacology, The University of Western Australia, Nedlands, Perth, Western Australia 6009, Australia.

§ To whom correspondence should be addressed.

FIG. 1 a, The activities of the mechanosensitive channel of large conductance (MscL; inverted triangle) and that of a smaller conductance (MscS; asterisk) in a representative patch excised from an *E. coli* envelope, in response to a ramp of suction (bottom diagram); b, MscL activities alone in a chromatographic fraction of the envelope upon reconstitution; and c, SDS-PAGE protein patterns of two simplified fractions, each from a series of chromatographic steps.

METHODS. Single-channel recording: For a, generation of giant spheroplasts, patch excision, and recording methods were all done essentially as described³. Pipette and bath solution: 200 mM KCl, 90 mM MgCl₂, 10 mM CaCl₂, 5 mM HEPES, pH 7.2. The patch was held at +20 mV pipette and examined at room temperature. The steady-state pipette pressure was measured with a manometer. The approximate course of the pressure ramp is shown. For b, an MscL-enriched fraction (see c) was reconstituted into azolectin liposomes (L-2-lecithin; type IIS, Sigma) at a 1/800 protein-to-lipid ratio, and the liposomes were sampled with patch-clamp pipettes as described⁵. Pipette and bath solution: 200 mM KCl, 40 mM MgCl₂, 5 mM HEPES, pH 7.2. Other parameters as in a. Chromatographic fractionation and patch-sampling (c): Individual or pools of fractions from each step were reconstituted into azolectin liposomes and 7–20 patches were sampled. Series 1: total *E. coli* membrane fraction was isolated and extracted with 2% octylglucoside (OG) as described⁵. OG extract (30 ml, 2 mg protein per ml) was loaded onto a DEAE-Sepharose CL-6B column in 2 mM EDTA, 2 mM EGTA, 1 mM DTT, 10 mM Tris buffer, pH 7.2. The majority of MscL activity was eluted with 300 mM NaCl and loaded onto a hydroxyapatite column, which was washed with 1 M MgCl₂ and then eluted with a linear phosphate gradient from 0 to 400 mM. Fractions 11–15 out of 25 were found active. They were pooled, concentrated and then passed through



a Superose 6 HR 10/30 FPLC gel-filtration column. Fractions 9–11 out of 15 were active and subsequently subjected to chromatofocusing in pH 7–4 gradient on a Pharmacia Mono-P HR 5/5 FPLC column. The flow-through fraction was highly active and had the simplest protein pattern on SDS-PAGE after staining with silver (c, left lane). Series 2: the OG extract was cut with ammonium sulphate (35% saturation). The supernatant was applied to a phenyl Sepharose CL-6B column and separated with a linear 40–0% gradient of ammonium sulphate. Fractions 17–23 (of 23 total) containing the MscL activity were pooled, concentrated and passed through a Superose 12 HR 10/30 FPLC gel-filtration column. Fractions 5–9 (out of 10) were active; fraction 7 had the simplest pattern (c, right lane). Fraction 7 was used in the reconstitution experiment in b.

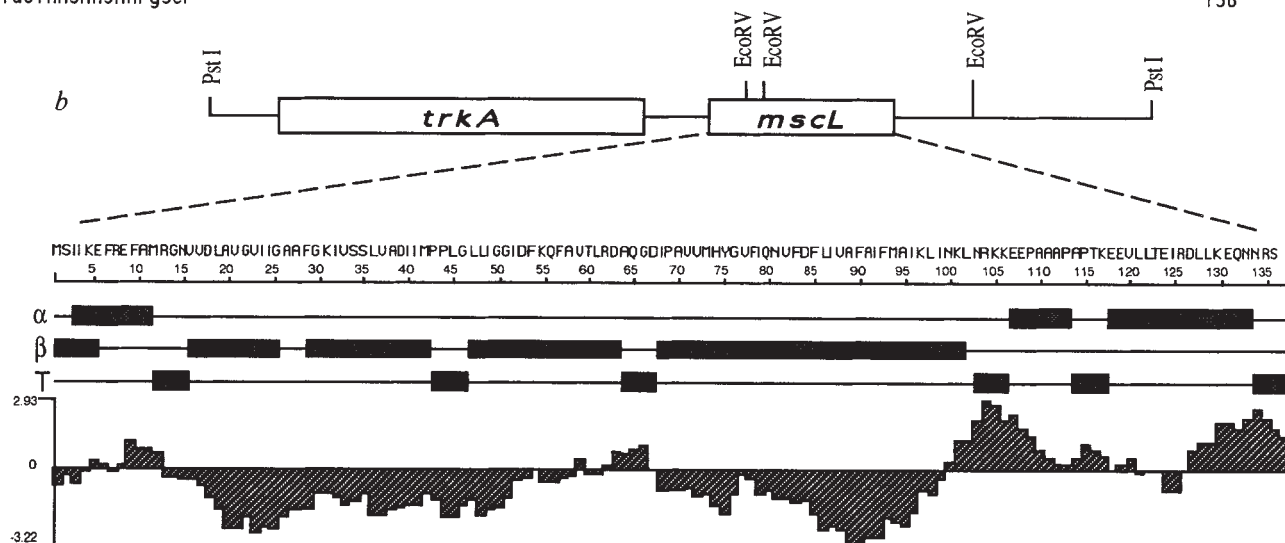
a
TATGGTTGTCGGCTTCATAGGGAGAAATACATGAGCATTATTAAAGAAATTCGCGAATTTGCGATGCGCGGGAACGTGGTGGATTTGGCGGTGGGTGTCATTATC 75
MetSerIleIleLysGluPheArgGluPheAlaMetArgGlyAsnValValAspLeuAlaValGlyValIleIle 25

GGTGGGCATTCGGGAGATTGCTCTTCACTGGTTGCCGATATCATCATGCCTCCTCTGGGCTTATTAATTGGCGGGATCGATTTTAACAGTTTGGCTGCACG 180
GlyAlaAlaPheGlyLysIleValSerSerLeuValAlaAspIleIleMetProProLeuGlyLeuLeuIleGlyGlyIleAspPheLysGlnPheAlaValThr 60

CTACGCGATGCGCAGGGGATATCCCTGCTGTTGTGATGCATTACGGTGCTTCATTCAAACGCTTTTGATTTTCTGATTGTGGCCTTTGCCATCTTTATGGCG 285
LeuArgAspAlaGlnGlyAspIleProAlaValValMetHisTyrGlyValPheIleGlnAsnValPheAspPheLeuIleValAlaPheAlaIlePheMetAla 95

ATTAGCTAATCAACAACTGAATCGGAAAAAGAGACACGAGCCGACCTGCACCACTAAGAGAGAGTATTACTGACAGAAATTCGTGATTGCTGAAA 390
IleLysLeuIleAsnLysLeuAsnArgLysLysGluGluProAlaAlaAlaProAlaProThrLysGluGluValLeuLeuThrGluIleArgAspLeuLeuLys 130

GAGCAGAATAACCGCTCTTAACAGCGCCTGAAGCAGAGACCAGTGGTAAAAAGTGATTACTTTCTTGCCACTGGCCTCCCGATTCCCCCGATTGCCATG 495
GluGlnAsnAsnArgSer * 136



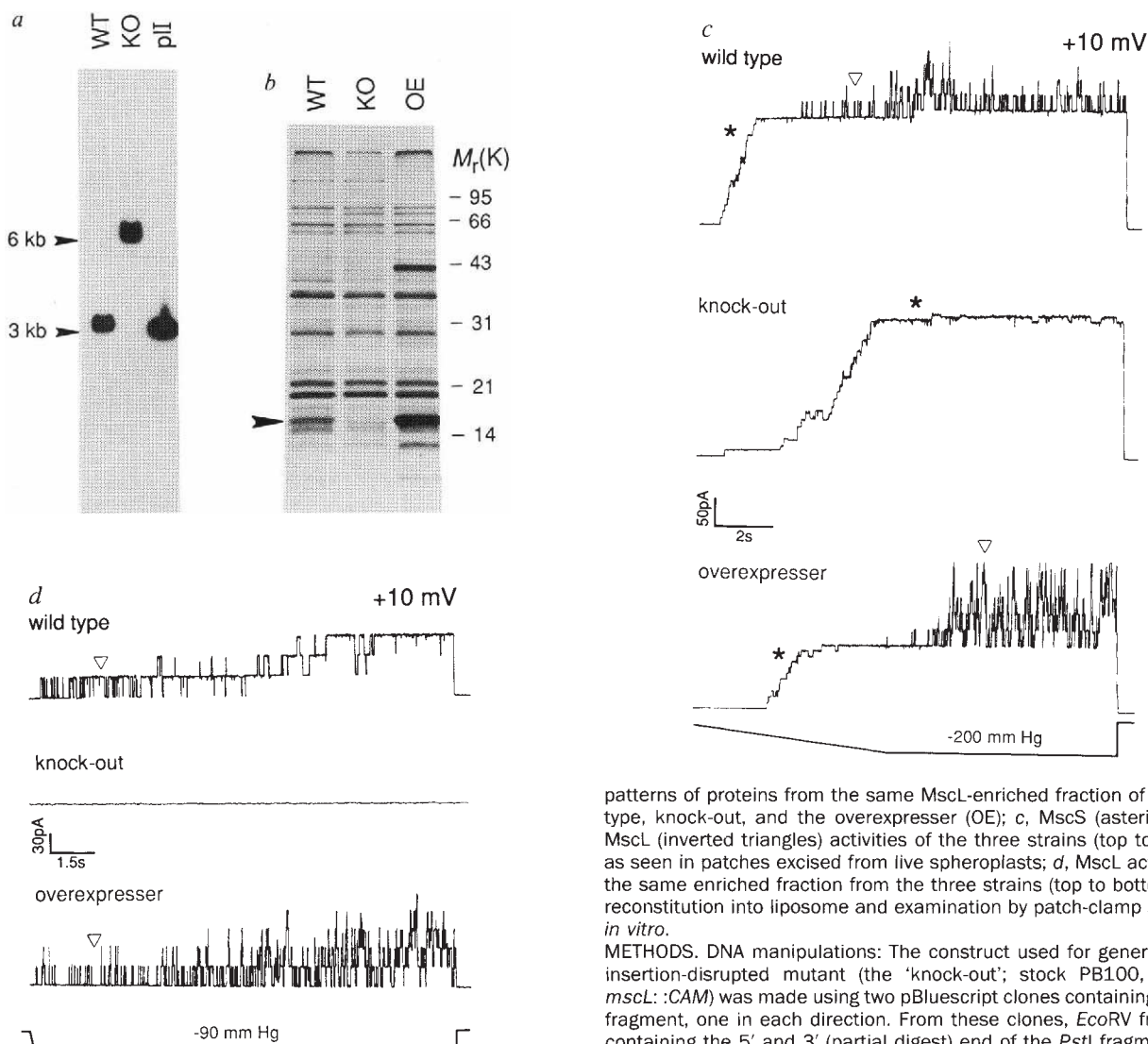


FIG. 3 Molecular and functional analyses of MscL in a parental wild-type strain, in a strain with its *mscL* disrupted by an insertion (knock-out), and a strain overexpressing a plasmid-borne *mscL* in a knock-out background (the overexpresser). **a**, Southern analysis of genomic DNA of the wild type (WT), the knock-out (KO), and the *mscL* ORF-containing subcloned *PstI* fragment (pII); **b**, SDS-polyacrylamide gel electrophoretic

patterns of proteins from the same MscL-enriched fraction of the wild-type, knock-out, and the overexpresser (OE); **c**, MscS (asterisks) and MscL (inverted triangles) activities of the three strains (top to bottom) as seen in patches excised from live spheroplasts; **d**, MscL activities of the same enriched fraction from the three strains (top to bottom) after reconstitution into liposome and examination by patch-clamp sampling *in vitro*.

METHODS. DNA manipulations: The construct used for generating the insertion-disrupted mutant (the 'knock-out'; stock PB100, AW405-*mscL*:*CAM*) was made using two pBluescript clones containing the *PstI* fragment, one in each direction. From these clones, *EcoRV* fragments containing the 5' and 3' (partial digest) end of the *PstI* fragment were isolated and sequentially subcloned into the *EcoRV* and *HincII* sites in pBluescript SK⁺ such that, in effect, part of *mscL* ORF (between the two *EcoRV* sites) was replaced with the part of the pBluescript multiple cloning site containing the *HindIII* consensus sequence. The Ω chloramphenicol (Ω Cm) gene (~3.4 kb) was then subcloned into the unique *HindIII* consensus site. The entire *PstI* fragment containing Ω Cm was then electroporated into the CAG732 *E. coli* strain, and bacteria were selected for Cm resistance, and tested for homologous recombination by PCR. Subsequently, the disrupted *mscL* gene was transduced into the AW405 with P1 by Cm selection. The expression vector p5-2-2 was generated by PCR using primers starting at nucleotide -14 to +425 with added linkers containing the *XhoI* consensus sequence. The fragment was digested and subcloned into the *XhoI* site of the expression vector pB10a. The insert was subsequently resequenced. PB100 was converted to *recA*⁻ by a *Tn10* transposition to generate PB101 (PB100-*srIA*:*Tn10*). PB101 was then transformed with p5-2-2 to form the overexpresser. The ~40K band in **b** is presumably encoded by pB10a and not by the insert. Standard techniques were used for Southern blot and hybridization. Total DNA extracted from each of the three strains was probed with the full-length *mscL* ORF at a stringency of 6 $\times$ SSC at 55 °C. Protein analysis: Wild-type AW405 bacteria, knock-out and overexpresser were each grown in 1 litre of LB to an absorbance of 0.8 at 595 nm. Membranes were isolated⁴, proteins extracted with octylglucoside and chromatographed on Q-Sepharose, followed by a 20% ammonium sulphate cut and then chromatography on phenyl Sepharose. Fractions collected at the end of 20 to 0% ammonium sulphate gradient were collected, concentrated and reconstituted into liposomes at 1:400 protein-to-lipid ratio, followed by patch-clamp analysis. Aliquots of each of the fractions were also precipitated with trichloroacetic acid and subjected to SDS-PAGE using equal protein loading per lane. Patch-clamp analysis of spheroplasts or liposomes was as described for Fig. 1.

◀ FIG. 2 **a**, Nucleotide and corresponding amino-acid sequence of the *mscL* gene; **b**, predicted secondary structure (α , α -helical; β , β -sheet; T, peptide turns) and a hydropathy profile of the MscL protein. Top diagram in **b** shows the position of *mscL* with respect to *trkA* and the location of restriction sites used in cloning. Negative numbers measure relative hydrophobicity in the plot. **METHODS.** N-terminal sequence was determined and sequencing primers were generated by the Biotechnology Center, Univ. Wisconsin. Gene cloning: A 3.1-kb *PstI* fragment from λ phage EC5-106 was subcloned into pBluescript SK⁺. Sequencing primers 19 nucleotides long, starting homology at -125 and +96 bases, were designed from information for this region in Genbank. For additional sequence, the main portion of the *mscL* ORF and its 5' flank were first subcloned as 415 bp and 500 (partial digest) *EcoRV* fragments into pBluescript, and then sequenced using Promega primers T3 and T7. To confirm the internal sequence, we used two additional primers, one for each direction, between +325 and +345. All sequence was obtained in both directions using a Sequenase kit (USB). Structural analysis was performed with DNA-Star software using the Chou-Fasman algorithm. The hydropathy plot used the Kyte-Doolittle algorithm with a window of 9 residues.

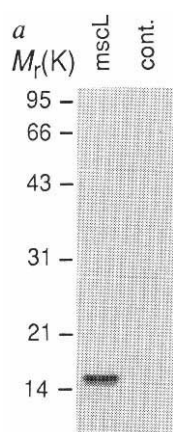
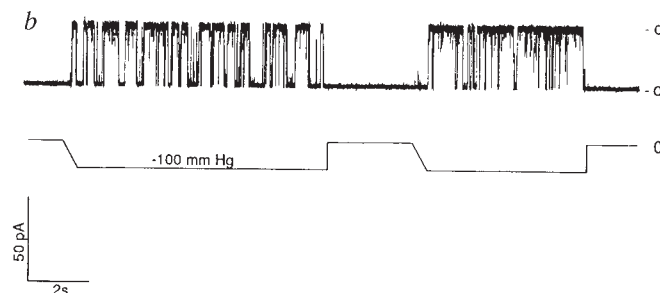


FIG. 4 Cell-free expression of the *E. coli* *mscL* gene. *a*, Autoradiogram of an SDS-PAGE gel of expression lysate driven by an *mscL*-containing (left lane) or control (right lane) vector. *b*, Activities of 2.5-nS mechanosensitive channels recorded at +20 mV in liposomes fused with a fraction from the expression lysate. Currents in the closed (c) or open (o) state are marked.

METHODS. A transcription-translation-coupled reticulocyte lysate system (Promega) was used. Canine microsomes were added as a neces-



sary supplement for membrane protein translation. The *mscL* was subcloned into the pSP64-polyA vector and used as template. After 2 h incubation of complete 25 μ l reaction mixture (for details, see Promega Technical Bulletin no. 126) with 35 S-methionine, microsomes were pelleted, washed and dissolved in 1% sodium deoxycholate. Proteins were precipitated with trichloroacetic acid, separated by SDS-PAGE, and the gel autoradiographed. For functional assay of mechanosensitive channel activity, 50- μ l reactions were set up with a full complement of unlabelled amino acids. Microsome fractions were washed, reconstituted into azolectin liposomes (2 mg lipid per preparation) and examined by patch-clamp as described for Fig. 1. In two independent experiments, 8 MscL channels were encountered in 20 patches for *mscL*-expressing lysate, and none was evident in 20 patches for the negative vector control.

We have disrupted the chromosomal *mscL* by inserting the chloramphenicol-resistance marker into the ORF. We have also re-introduced into this 'knock-out' background the wild-type *mscL* gene borne on a plasmid under a strong and inducible promoter (Fig. 3a). Examination of membrane patches from live giant spheroplasts derived from the wild-type parent showed 4–10 MscL conducting units (namely, 3-nS conductance appearing upon suction) per patch in all 10 patches examined, no MscL units in any of the 50 patches from the knock-out mutant, but more than 40 units per patch in every one of the 6 patches excised from the rescued strain after induction of the expression plasmid (Fig. 3c). Consistent with the electrophysiological findings, protein patterns of the same chromatographically enriched fraction from these three strains showed that the 17K protein was missing in the knock-out mutant and was enriched in the rescued and overexpressing strain (Fig. 3b, arrowhead). Such fractions were individually reconstituted into liposomes and examined by patch clamp. We encountered 5 ± 2 units of 3-nS conductance per patch (mean $\pm$ s.d., $n=6$) excised from liposomes reconstituted with the wild-type fraction, no such conductance in any of the 11 patches from liposomes with the knock-out fraction, and 37 ± 9 ($n=5$) from those with the 'over-expression' fraction (Fig. 3d).

When an *mscL*-containing vector was used to drive a reticulocyte-lysate expression system, a 17K protein was synthesized (Fig. 4a). MscL mechanosensitive channel activity was observed in liposomes incorporating the expressed material (Fig. 4b). Empty vector did not drive protein expression or generate chan-

nel activity. Thus, *mscL* alone is not only necessary but sufficient for the entire structure responsible for ion conduction and mechanosensitivity.

Previous studies have suggested that mechanosensitive channels in *E. coli* function in detection of and/or adaptation to different osmotic changes. Interestingly, *mscL* is only 133 nucleotides downstream from *trkA* and both genes are expected to be transcribed clockwise. *trkA* encodes a protein for the uptake of K^+ , the most important osmoticum, and the TrkA-dependent K^+ -uptake activity is regulated by the osmolarity of the medium¹¹. In addition, it has been shown¹² that *E. coli* cells, when put in distilled water, jettison their small solutes rapidly through certain 'holes' while maintaining their macromolecules and viability. We and others^{5,13} have proposed that some of the mechanosensitive channels could be these 'holes'. MscL may serve as a bleed of osmotica at critical pressures. Indeed, comparisons of the *mscL* knock-out and the wild-type strains have shown dramatic differences in hypotonic-shock-induced membrane permeability for relatively large osmolytes, such as ATP (S.I.S., unpublished results). Although we found that insertional disruption of *mscL* is not lethal in laboratory cultures, MscL and MscS may be redundant in functions called upon only under transient acute osmotic stress and are not required for vegetative growth. In any event, these small channel proteins that appear to respond directly to stretch forces in the lipid bilayer may serve as a model to study the molecular basis of mechanosensation and a means to study aspects of bacterial physiology. □

Received 1 November 1993; accepted 21 January 1994.

1. Sachs, F. in *Sensory Transduction* (eds Corey, P.D. & Roper, S. D.) 241–260 (Rockefeller Univ. Press, New York, 1992).
2. Martinac, B. in *Thermodynamics of Membrane Receptors and Channels* (ed. Jackson, M. B.) 327–352 (CRC, Boca Raton, 1993).
3. Martinac, B., Buechner, M., Delcour, A. H., Adler, J. & Kung, C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2297–2301 (1987).
4. Delcour, A. H., Martinac, B., Adler, J. & Kung, C. *Biophys. J.* **56**, 631–636 (1989).
5. Sukharev, S. I., Martinac, B., Arshavsky, V. Y. & Kung, C. *Biophys. J.* **65**, 177–183 (1993).
6. Martinac, B., Adler, J. & Kung, C. *Nature* **348**, 261–263 (1990).
7. Hamann, A., Bossemeyer, D. & Bakker, E. P. *J. Bact.* **169**, 3138–3145 (1987).
8. Daniel D. L. & Blattner, F. R. *Nature* **325**, 831–832 (1987).
9. Cowan, S. W. *et al. Nature* **358**, 727–733 (1992).

10. Blachly-Dyson, E., Peng, S., Colombini, M. & Forte, M. *Science* **247**, 1233–1236 (1990).
11. Epstein, W. *FEMS Microbiol. Rev.* **39**, 73–78 (1986).
12. Britten, R. J. & McClure, F. T. *Bact. Rev.* **26**, 292–335 (1962).
13. Bernier, C., Coulombe, A., Scabo, I., Zoratti, M. & Ghazi, A. *Eur. J. Biochem.* **206**, 559–565 (1992).

ACKNOWLEDGEMENTS. We thank J. Mecas, J. Rossmeissl, C. Gross, V. Arshavsky and R. Guy for advice; J. Adler and C. Gross for bacterial stocks; and T. Yura for plasmid pB10a. The λ clone set was developed by D. L. Daniels. F.R.B. is supported by a grant from the NIH; S.I.S. is a trainee in the DOE/NSF/USDA Collaborative Program on Research in Plant Biology Program; P.B. was a trainee in the NIH Postdoctoral Training Program in Medical Genetics and is now a DOE Energy Bioscience Research Fellow of the Life Sciences Research Foundation.

Multi-target PCR analysis by capillary electrophoresis and laser-induced fluorescence

Wei Lu, Dai-Shu Han, Ju Yuan and Jean-Marie Andrieu

Quantitative analysis of polymerase chain reaction (PCR) amplified HIV-1 DNA or cDNA fragments is attained using an automated system that combines capillary-gel electrophoresis (CGE) for high-efficiency separation and laser-induced fluorescence (LIF) for high-sensitivity detection. This system enables the detection of PCR-amplified multiple target DNA or cDNA in the same tube by a single injection with high precision.

In many applications in molecular biology and biomedical research, quantitative detection of target sequences in native nucleic acids is frequently required. Although the polymerase chain reaction (PCR) technology has enabled significant advances in the identification and quantitation of pathogenetic or physiological nucleotide sequences^{1,2}, the complexity of conventional methodologies (such as radioactive or non-radioactive hybridization) used for post-PCR analysis has severely limited the application of DNA probe technology in molecular diagnosis. In addition, the sensitivity and linearity of quantitative PCR can be influenced by pipetting errors or sample loss during the arduous hybridization procedure. The ultimate goal for optimizing a quantitative assay, therefore, is to explore the feasibility of using a direct and automated post-PCR detection system.

Capillary-gel electrophoresis (CGE) is an elegant combination of traditional electrophoresis and modern liquid chromatography that employs an anticonvective medium such as polyacrylamide in narrow-bore (internal diameter (i.d.) 50–100 µm) capillaries to generate fast, high-efficiency separations of nucleic acid molecules up to unit base resolution³. The availability of automated systems combining CGE and laser-induced fluorescence (LIF) detection (LIF-P/ACE, Beckman Instruments, Fullerton, California)⁴ provides an easy way to evaluate the potential application of this technology. In order to demonstrate the usefulness of LIF-CGE for direct analysis of target nucleic acid sequences in the PCR reaction tube, this technique has been used to analyse HIV-1 DNA or complementary DNA (cDNA) fragments amplified by PCR or a reverse transcription (RT)-linked PCR assay⁵. This system has also been compared to conventional ³²P-labelled hybridization for quantitating HIV-1 virion concentration in sera taken from 30 individuals infected with HIV-1 (stage I–IV, Centers for Disease Control and Prevention, Atlanta, Georgia).

Detection

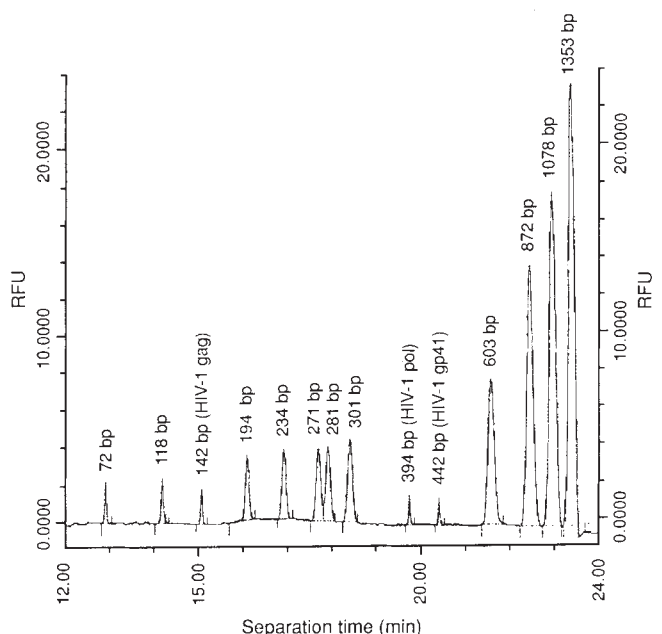
Using a 10-copy HIV-1 DNA template (HIV-1 pBH10-R3) over a background of NATURE · VOL 368 · 17 MARCH 1994

1 µg DNA extracted from human peripheral blood mononuclear cells (PBMC), multi-target PCR yields of 142, 394 and 442 base pairs (bp) were obtained by amplification with HIV-1 *gag* (p24), *pol* and *env* (gp41)-specific oligonucleotide primers (see legend to Fig. 1). The amplified multiple DNA fragments were detected in a single injection by LIF-CGE (Fig. 1). The peaks of amplified sequences were identified by the retention time, using the calibration curve derived from the simultaneous injection of molecular weight standards (*Hae* III restric-

tion digest of ϕ -X 174 RF DNA). Specificity was confirmed by Southern blot hybridization (data not shown). The concentration of PCR product was determined by the peak height on a CGE chromatogram, using calibration factors generated from serial dilutions of known concentrations of a DNA standard. The precision (s.d./mean $\times 100$ per cent) for replicate injection reached 1 per cent. In principle, the detection limit of the LIF detector is about 10^{-18} mole (attomole), which is comparable to that of radioactive detection. Thus, the multi-target PCR yields

FIG. 1 Detection of multi-target PCR-amplified HIV-1 *gag*, *pol*, *env* sequences by automated laser-induced fluorescence (LIF)-linked capillary-gel electrophoresis (CGE). A concentration of HIV-1 DNA templates was obtained by diluting plasmid DNA (HIV-1 pBH10-R3) in background DNA from human peripheral blood mononuclear cells (PBMC). A 10-µl aliquot of plasmid-based amplicon containing 10 HIV-1 DNA templates and 1 µg PBMC DNA was added to 90 µl of PCR reaction mixture containing three units of AmpliTaq DNA polymerase and 100 pmol of each HIV-1 primer pair specific for *gag* (p24) gene

(SK145/SK431), *pol* gene (GGAGAGCAATGGCTAGTGA/TACTCCTTGACTTTGGGG), and *env* (gp41) gene (CTGTTGCAACTCACAGTCTG/TATCCCTGCCTAACTCTAT). PCR was run at 94 °C for 5 min, followed by 40 cycles (94 °C for 60 s, 56 °C for 60 s and 72 °C for 90 s). A 20-µl aliquot of amplified product was mixed with a 10-µl aliquot of molecular weight standards (*Hae* III restriction digest of ϕ -X 174 RF DNA) in a 50-µl mini-vial (Beckman Instruments) and placed in the automated LIF-CGE (P/ACE, Beckman) for post-PCR detection. The injection was 90 s at 5 kV. Separation was run at 7.4 kV for 24 min in a capillary (37 cm length, 100-µm i.d.) filled with an electrophoresis gel buffer containing DNA-fluorescence stain molecules (LIFluor dsDNA 1000 Kit, Beckman). The chromatograms were visualized and analysed automatically using GOLD Software (Beckman). The peak high was expressed in relative fluorescence units (RFUs) and the retention time was given in minutes. The molecular weight standards consisting of 11 DNA fragments and PCR-amplified fragments (HIV *gag*, *pol* and *env* genes) were distinguished by the retention time; the relative quantity of each DNA fragment was represented by the peak high.



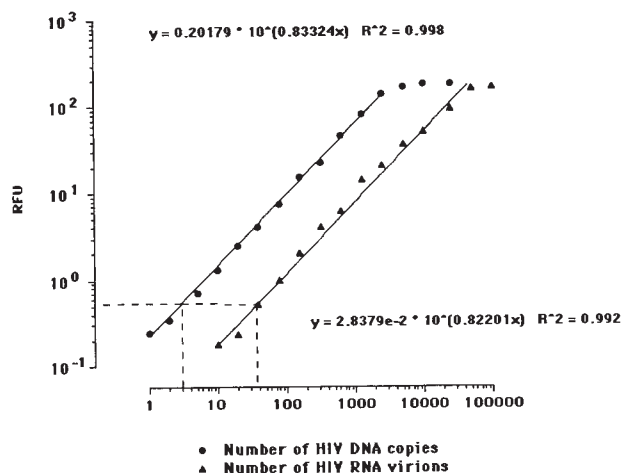


FIG. 2 Linearity of LIF-CGE analysis of quantitative PCR or RT-PCR products. Serial dilutions of HIV-1 DNA templates ranging from 1 to 25,000 copies were subjected to a 40-cycle PCR with *gag* primers SK145/431 as described in the legend to Fig. 1. RNA extracted from serial dilutions of HIV-1 virions (ranging from 10 to 100,000 viral particles) was reverse-transcribed with 20 pmol of 3' primer SK431 as described elsewhere⁵. The resulting cDNA was then amplified for *gag* gene as described above. The PCR or RT-PCR product of each dilution was analysed as described in Fig. 1.

(2–5 femtomole) of the 10-copy target DNA template were readily detected by LIF-CGE (Fig. 1).

Quantification of target DNA/RNA

In practice, the unknown concentration of target DNA or RNA can be estimated by measuring the concentration of PCR-amplified DNA or RT-PCR-amplified RNA (expressed in any units of measurement), using a standard curve generated by coamplification (either internally or externally) of known concentrations of reference nucleic acids². By limiting dilution of HIV-1 DNA in PBMC, a 142-bp PCR product of HIV-1 *gag* genome amplified from a single template over a background of 1 µg DNA (10^5 PBMC) was identified using LIF-CGE (data not shown). Such a concentration of one template per tube fitted the pattern of Poisson distribution model where 63 per cent of replicate tubes have one or more target sequence. The peak height of the amplified fragment increased proportionally with the concentration of target DNA. A linear range of over three orders of magnitude of PCR or RT-PCR yields was demonstrated (Fig. 2).

Measuring virion concentration

Using a previously validated quantitative RT-PCR assay^{5,6}, the efficiency of the post-PCR detection methods of LIF-CGE and radioisotopic procedures were compared for measuring the concentration of HIV-1 RNA virions in sera of 30 infected patients (seven cases with primary infection, 13 who were asymptomatic, and 10 patients with AIDS). Briefly, HIV-1 virion RNA was directly extracted from a 500-µl aliquot of patient sera using an RNA isolation kit (RNAzol, WAK-Chemie Medical, Bad Homburg, Ger-

many). After adding 1 µg tRNA, the RNA sample was pelleted and resuspended in 50 µl of RNA-storing solution. A 10-µl aliquot of each RNA sample (containing the RNA of 100-µl serum) was added to each (90 µl) of RT-PCR reaction mixture containing 10 units of recombinant Moloney murine leukaemia virus reverse transcriptase (Boehringer Mannheim, Mannheim, Germany) and three units of *Taq* DNA polymerase (AmpliTaq DNA polymerase, Perkin-Elmer, Norwalk, Connecticut) for HIV-1 *gag* gene amplification. RT-PCR was run for 25 min at 42 °C, followed by an additional 5 min at 94 °C and then 40 cycles of PCR (see legend to Fig. 1). The resulting RT-PCR

product was quantitated by both automated LIF-CGE (20-µl aliquot) and DNA hybridization (10-µl aliquot) assays. Noninfectious (heat-inactivated) HIV-1 stock of known virion concentration was used as an external standard against which the serum virion concentration was determined using a formula derived from the standard curve for each batch of the RT-PCR assay⁶. Each serum sample was tested in three different batches and each batch was done in triplicate. As summarized in the table, there was concordance between the concentration of serum virions measured by LIF-CGE and that measured by DNA hybridization ($P=0.493$). The precision of the serum virion quantitation assay was significantly better using the LIF-CGE system than with the radioactive assay ($P<0.001$), representing a more than fivefold improvement in both intratest and intertest variation.

Discussion

The LIF-CGE data show that different target sequences amplified by multi-target PCR can be identified simultaneously by a single injection. This approach offers considerable advantages over other methods: it simplifies the experimental procedure (direct detection can be accomplished within 30 min), and PCR products can be seen automatically. The linearity between template DNA or RNA dilution and PCR or RT-PCR products de-

tected by LIF-CGE demonstrates that LIF-CGE is an accurate technique for analysing amplified fragments quantitatively. Agreement in the quantitative PCR or RT-PCR results as determined by LIF-CGE and DNA hybridization indicates that LIF-CGE is likely to be a useful tool for quantitating specifically amplified target sequences. The use of LIF-CGE (leading to high precision for detection) overcomes the variation resulting from conventional post-PCR experiments, and thus enhances the ability of PCR or RT-PCR to quantitate any sequences of interest in native nucleic acids.

The reduction in intratest and intertest variation in RT-PCR quantitation by LIF-CGE (see table) further strengthens the credibility of simply using an external virion standard for HIV-1 RNA quantitation. Unlike competitive RT-PCR, which applies an *in vitro* constructed RNA (distinct from native RNA structure) as internal standard⁷, noncompetitive RT-PCR using HIV-1 virion itself as external standard is much less cumbersome and eliminates the discord in RT efficiency resulting from distinct genetic structure between the tested RNA and the calibrated RNA. In addition, due to natural protection of viral envelope the virion RNA is stable during viral stock preparation and storage⁶, and thus can be easily used as a constant reference in different laboratories. Serum RT-PCR data reveal that the LIF-CGE detection system provides a practical means of quantitating HIV-1 virion concentration in the sera of infected patients. Most (90 per cent) patient sera has a virion concentration ranging from 200 to 500,000 viral particles per millilitre (within the detection range of the assay system) and can be

Comparison of CGE-LIF detection and DNA hybridization (Southern blot) for measuring HIV-1 virion concentrations in sera of individuals infected with HIV-1 at different stages of the infection

	Virion concentration ($\times 10^4$ ml ⁻¹) as determined by:	
	CGE-LIF	Southern blot
Primary infection (CDC* stage I, n=7)		
Geometric mean	18	14
Range	0.8–1,170	0.9–959
Asymptomatics (CDC stage II/III, n=13)		
Geometric mean	17	19
Range	0.05–1,097	0.02–813
ARC/AIDS (CDC stage IV a–d, n=10)		
Geometric mean	43	36
Range	9–444	6–226
Intratest variation (n=30)†		
Mean $\pm$ s.d.	7% $\pm$ 4%	32% $\pm$ 7%
Intertest variation (n=30)†		
Mean $\pm$ s.d.	10% $\pm$ 6%	51% $\pm$ 11%

* Centers for Disease Control and Prevention (CDC)

† Intratest variation expressed as the mean ($\times 100\%$) of triplicate tests

‡ Intertest variation expressed as the mean ($\times 100\%$) of three discrete measurements of each serum

measured by a single test. Only a few cases (less than 10 per cent) have a serum virion concentration over 500,000 virus particles per millilitre (exceeding the detection limit of the assay system) and need to re-run the RT-PCR using a tenfold less quantity of RNA as template. The RT-PCR assay (detection limit, 10 virions) showing 100 per cent sensitivity for HIV virion detection in the patients' sera (table, refs 7–9), represents a gain of at least four orders of magnitude compared to the classical HIV p24 antigen enzyme-linked immunosorbent assay (ELISA), the detection threshold of which is 10 pg (corresponding to 10^5 virions). The duration of the whole process (serum RNA extraction, RT-PCR and LIF-CGE detection) is accomplished within 4 hours.

The potential for contamination through carry-over appears to be another limitation in the use of conventional PCR for molecular diagnosis, and numerous preventive strategies have been developed to minimize the probability of this occurring^{10–12}. Multi-target-PCR coupled with LIF-CGE allows the screening of several fragments of a target genome in the same reaction tube and the resulting LIF-CGE profile would be helpful in discerning a false-positive PCR result.

High-sensitivity detection of retroviral genomes is of considerable importance in the management of HIV transmission. Fol-

lowing HIV infection, an individual generally remains seronegative for a period of time that can range from 10 days up to 6 months (referred to as the 'window period')^{13,14}. During this period of time, due to the limited sensitivity of ELISA (see above), the serum HIV p24 antigen is often negative^{14,15}. It has been reported that even though donated organs and tissues are screened before transplantation by anti-HIV antibody and p24 antigen tests, HIV transmission during the window period has occurred¹⁴. Thus, direct screening of viral nucleotides by PCR/RT-PCR assays would abrogate this window phase and should be more effective than the antigen-antibody immunosorbent assays for the overall control of HIV transmission. It will, for example, discount optimally the risk of transfusing blood or blood products derived from seronegative but HIV-infected donors.

In conclusion, these data demonstrate the significant benefit that automated LIF-CGE has in the identification and quantitation of target nucleotide sequences. Apart from its significant potential in the area of molecular diagnosis, this technology should have wide application in the monitoring of antiviral therapy and the analysis of multiple gene expression within a homogenous cell population. □

Wei Lu, Dai-Shu Han, Ju Yuan and Jean-Marie Andrieu are at Laboratory of Tumor Immunology, Hopital Laennec, Université de Paris V, Paris 75007, France. This study was supported by Agence Nationale de Recherche sur le SIDA, Association de Recherche pour le Traitement des Seropositifs, and Association pour la Recherche et l'Etude des Maladies du Sang. For more information, fill in reader service number 100.

1. Erlich, H. A., Gelfand, D. & Sninsky, J. J. *Science* **252**, 1643–1656 (1991).
2. Andrieu, J. M. in *Viral Quantitation in HIV Infection* 191 (John Libbey Eurotext, Paris, 1991).
3. Luckey, J. A. et al. *Nucleic Acids Res.* **18**, 4417–4421 (1990).
4. Albin, M., Weinberger, R., Sapp, E. & Moring, S. *Analyt. Chem.* **63**, 417–422 (1991).
5. Lu, W. & Andrieu, J. M. *J. infect. Dis.* **167**, 1014–1020 (1993).
6. Lu, W. & Andrieu, J. M. *J. infect. Dis.* **167**, 1498–1499 (1993).
7. Platack, M. et al. *Science* **259**, 1749–1754 (1993).
8. Lu, W., Eme, D. & Andrieu, J. M. *Lancet* **341**, 113 (1993).
9. Lu, W., Britton, S. & Andrieu, J. M. *AIDS* **7**, (Suppl. 2), S111–S112 (1993).
10. Isaacs, S. T., Tessman, J. W., Metchette, K. C., Hearst, J. E. & Cimino, G. D. *Nucleic Acids Res.* **19**, 109–116 (1991).
11. Ou, C. Y. & Schochetman, G. *BioTechniques* **10**, 442–445 (1991).
12. Shuldiner, A. R., Tanner, K., Moore, C. A. & Roth, J. *BioTechniques* **11**, 760–763 (1991).
13. Horsburgh, C. R. et al. *Lancet* **334**, 637–640 (1989).
14. Simonds, R. J. *AIDS* **7**, (Suppl. 2), S35–S38 (1993).
15. Lefrere, J. J., Courouce, A. M., Rouger, P., Duedari, N. & Elghouzi, M. H. *Lancet* **339**, 999–1000 (1992).

Tools of the trade

For the immunologist — an *in vitro* system for studying leukocyte transendothelial migration, assays for free radical markers, antibodies to nitric oxide synthases and an assortment of microplates.

COLLABORATIVE Biomedical Products/Becton Dickinson has introduced the BioCoat Leukocyte Traffic Environment as an *in vitro* system for the study of **leukocyte transendothelial migration** (Reader Service No. 101). This system is said to provide an *in vitro* model of the luminal surface of blood vessels to study leukocyte extravasation in a physiological and reproducible manner. The BioCoat Leukocyte Traffic Environment integrates specially formulated

low-serum medium supplemented with culture supplements, BioCoat Fibronectin Cell Culture Inserts (containing 3 µm pore size membranes) and a control chemotactic agent. The system is said to promote the establishment of confluent endothelial monolayers with barrier function within two days. Stated applications include studies of migration and adhesion of several leukocyte types, studies of the effects of anti-inflammatory drugs on leukocyte traffic, and the separation and characterization of migrating and nonmigrating leukocytes. Each environment contains E-STIM endothelial cell culture medium, endothelial cell growth supplement, EGF, IL-1β, and 24 BioCoat fibronectin cell culture inserts (3 µm pore size).

Incstar has increased its **custom immunization services** through its Atlantic Antibodies farm facility in Windham, Maine (Reader Service No. 102). Since the acquisition of Atlantic Antibodies in 1987, Incstar says that it has constructed 18,000 square feet of new animal facilities. The 120-acre farm currently has the capacity to house over 700 animals for polyclonal antisera production. Host species include chickens, don-



Custom immunization services from Incstar's Atlantic Antibodies farm facility.

keys, goats, guinea pigs, rabbits and sheep. The plasmapheresis capabilities of Atlantic Antibodies have also been expanded. The farm facility provides bulk antisera, calibrators and controls for immuno-assay manufacturers, as well as offering services to clients with their own immunogens. A custom immunization service booklet outlining the services provided by Incstar's Atlantic Antibodies is now available on request.

■ Going with the flow

Exalpa has introduced a new *Quick Reference Guide for flow cytometry reagents* (Reader Service No. 103). The company



***In vitro* system for studying leukocyte transendothelial migration.**

PRODUCT REVIEW

also offers a more convenient package size of 100 tests per vial for single-colour and 50 tests per vial for two-colour reagents. Each reagent is offered in a purified form or conjugated with fluorescein isothiocyanate, biotin or phycoerythrin. Exalpa's line of reagents should be of interest to researchers studying AIDS, T and B cell subsets, NK cells, T-cell activation and cellular adhesion.

Becton Dickinson Immunocytometry Systems has introduced a new family of second-step **conjugates for multi-colour flow cytometry** that are intended for research in the fields of immunology, microbiology, oncology, developmental biology and

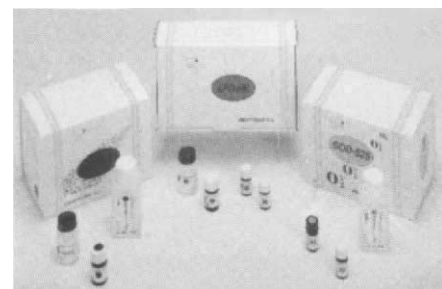


Anti-mouse, isotype-specific, second-step conjugates from Becton Dickinson.

stem cell maturation and development (*Reader Service No. 104*). The four new rat-anti-mouse, isotype-specific, second-step conjugates are RAM IgG₁ PE, RAM IgG_{2a+b} PE, RAM IgG₁ PerCP and RAM IgG_{2a+b} PerCP. These new anti-mouse conjugates may be combined with appropriate primary antibodies and conjugates to produce a wide variety of one-, two- and three-colour, single-laser analyses. They also eliminate the need to perform biotin and fluorochrome conjugations. Becton Dickinson says that because PerCP has a sharp emission profile, which peaks at about 680 nm, virtually no compensation with PE is required when used in multi-colour experiments that include FITC and PE.

■ Assays

R & D Systems recently launched a range of **assays for free radical markers** that are intended as a means of standardizing the methodology of free radical research in the fields of AIDS, inflammation and aging (*Reader Service No. 105*). The new products



Assays for free radical markers.

include colorimetric assays for glutathione, lipid peroxidation and lactoferrin, and immunoassays for human myeloperoxidase, human lactoferrin and human plasma selenium glutathione peroxidase.

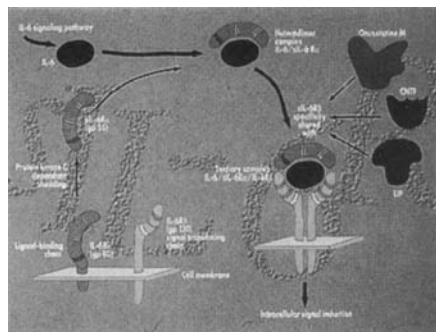
The Ridascreen Mycoplasma IFA is a new immunofluorescence test available from r-biopharm that is designed for the **detection of mycoplasma in cell cultures** (*Reader*



Mycoplasma detection kit.

Service No. 106). The kit contains a fluorescein-conjugated monoclonal antibody, a fluorescein-conjugated secondary antibody for enhancement of the signal (which is optional), mounting fluid and two positive and negative control slides. The kit contains sufficient reagents for testing 50 samples and the reagents have a shelf-life of up to a year when stored at 2–8 °C.

AMAC's soluble IL-6 receptor (sIL-6R) assay, a one-step 'sandwich' enzyme ligand immunoassay, can be used to **measure sIL-6R in serum, plasma and culture media** (*Reader Service No. 107*). The microtitre



Soluble IL-6 receptor assay measures only biologically active sIL-6R.

plate wells are coated with anti-IL-6 antibody. The standards, samples and conjugate (containing IL-6 and anti-sIL-6R antibody conjugated to alkaline phosphatase) are added to the plate and incubated for 16–20 h at room temperature on a shaker. The sIL-6 in the sample is bound to the wall of the well via its ligand, IL-6. This property makes it possible to measure biologically active sIL-6R. After washing the plate, a chromogenic substance is added and the plate incubated at room temperature for a further 30 min on a shaker. The plate is then read on a spectrophotometer at 405 to 415 nm. The assay is supplied in a 96-test kit format containing all the reagents necessary to perform the assay.

New from PerSeptive Diagnostics is the TiterZyme interleukin-1 β (IL-1 β) 'sandwich' enzyme immunoassay kit for the quantitative **determination of mouse IL-1 β levels** in serum or cell culture supernatant (*Reader Service No. 108*). Wells are coated with mouse IL-1 β antibody to provide consistency and flexibility, and standards are supplied prediluted and lyophilized. Results can be obtained in 3 hours and the stated sensitivity is 15.4 pg ml⁻¹ in buffer and 20.7 pg ml⁻¹ in serum diluent. An optional, room temperature 'shaking' format is said to increase sensitivity to 8.4 pg ml⁻¹ and 4.71 pg ml⁻¹ in buffer and serum diluent, respectively.

Biolab's **EBNA IgG ELISA test** uses a purified recombinant EBNA-1 antigen from which the glycine-alanine copolymer region has been deleted (*Reader Service No. 109*). This is said to eliminate crossreactivity due to antibodies present in the serum from patients with infectious mononucleosis or rheumatoid arthritis or due to normal cellular proteins. The EBNA IgG ELISA test is one of a new series of colour-controlled ELISA tests for infectious diseases in which the test procedure and the interpretation of the results have been simplified and standardized. The EBNA IgG ELISA test features multiple colour changes that mark each important step in the procedure; this is intended to allow close monitoring of the specimen and the addition of reagents, while, at the same time, reducing procedural errors. Combined into an EBV serological profile with ELISA tests for VCA-IgG and IgM, the EBNA IgG ELISA should provide useful information on EBV-associated disease.

■ Antibodies

Transduction Laboratories introduces **antibodies to several nitric oxide synthases (NOS)**, cell-type-specific enzymes that catalyse the synthesis of nitric oxide (NO)

Immunofluorescence microscopy using macNOS monoclonal antibody on a mouse macrophage cell line treated with γ -interferon and LPS.



(*Reader Service No. 110*). Monoclonal antibodies to the brain-specific, endothelial cell-specific and the inducible NOS are available. Quality control data are provided for western blot, immunoprecipitation and immunofluorescence analyses.

The proto-oncogene *c-kit* from Kamiya Biomedical encodes a tyrosine kinase receptor to stem cell factor (SCF), also known as mast cell growth factor, *kit*-ligand and steel factor (*Reader Service No. 111*). *C-kit* and SCF are important in the development of haematopoietic cells, germ cells and melanocytes. **Anti-mouse *c-kit* monoclonal**

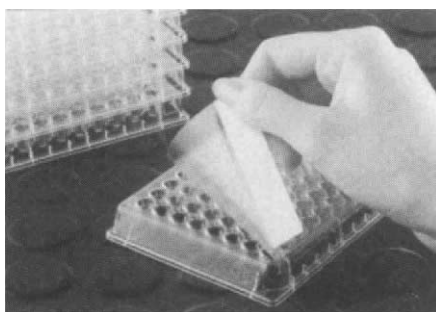
antibody ACK2 is specific to the extracellular domain of the proto-oncogene *c-kit*. ACK2 inhibits the action of SCF both *in vitro* and *in vivo*. When injected into mice, anti-mouse *c-kit* ACK2 causes the elimination of all haematopoietic cells. In pregnant mice, ACK2 causes alteration in coat colour in the progeny due to the blocking of melanocyte development within the embryo. Kamiya Biomedical says that in adult male mice, an intravenous injection of ACK2 results in a depletion in the differentiating type A spermatogonia. Anti-mouse *c-kit* ACK2 can be used for IHC staining (paraffin-embedded and frozen tissue sections), flow cytometry, western blotting, and *in vivo* and *in vitro* blocking of SCF-dependent activation.

Monoclonal **antibodies to cell-surface antigens and signal transducers**, manufactured by Medical and Biological Laboratories, are now available through PanVera (Reader Service No. 112). Highlighting the product line is the anti-Fas monoclonal antibody that reacts specifically with human Fas antigen and can induce apoptosis in cells expressing the human Fas antigen. It also shows utility in flow cytometry and immunoblotting for the study of immunology, embryology, AIDS and cancer. The nonradioisotopic protein kinase system is intended for measuring the activities of protein kinase C and cAMP-dependent protein kinase, using an anti-phosphoserine monoclonal antibody. The anti-phosphoserine monoclonal is said to show specificity for the phosphorylated form of human, bovine and porcine glial fibrillary acidic protein (GFAP) and negligible reactivity towards the dephosphorylated forms. After incubation of a protein kinase with nonradioactive ATP in 96-well plates coated with synthetic G1 peptide, -Arg-Arg-Arg-Val-Thr-Ser-Ala-Ala-Arg-Arg-Ser-(residues 3-13 of GFAP), the phosphorylated product is detected by using the mouse anti-phosphoserine monoclonal antibody and peroxidase-labelled goat anti-mouse IgG.

■ Handy tools

Crystal/Mount is a new **aqueous mounting medium** available from Biogenesis that is said to permit the long-term storage of antigens localized with chromogens that are soluble in organic solvents (Reader Service No. 113). Such chromogens include AEC, BCIP/INT, BCP/NBT and naphthol As-MX phosphate/Fast Red TR salt. The mounting medium can also be used for fluorescent-stained tissue sections where it is said to have a fluorescence stabilizing effect. It is a permanent medium drying to a glasslike finish over the section, which allows the slide to be filed in standard storage systems.

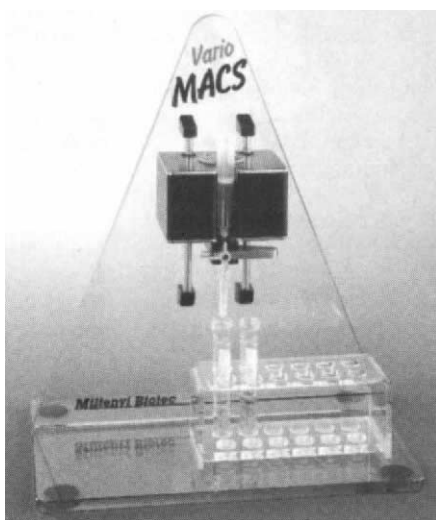
Elkay's **Sealplate microplate adhesive film** for 6- to 96-well plates is designed to prevent sample loss or evaporation in wells during incubation, transport, agitation or



Sealplate adhesive film for microplates.

storage (Reader Service No. 114). Perforated end tabs make sealing easier and more secure. The film (as well as its acroolefin adhesive) is nontoxic and inert, making it compatible with most procedures performed in microplates. Sealplate film is available in polyester material thin enough to allow penetration by probes or pipette tips, or in heat-resistant, thick polypropylene material for applications requiring temperatures up to 125 °C such as thermal cycling. The films are designed to be impermeable to moisture or oxygen. The acrylic base of the adhesive is said to resist most petroleum solvents and lower alcohols. Sealplate film is available in sterile or nonsterile forms, packaged 100 films to a box.

New from Miltenyi Biotec is the **VarioMACS magnetic cell separation system**, which combines the benefits of the original MAC's colloidal microbeads with the cell isolation capability of MiniMACS (Reader Service No. 115). The system, which



VarioMACS magnetic cell separator.

features a height-adjustable magnet and is intended for depleting large numbers of unwanted cells followed by enrichment of the target population, uses the complete line of MACS microbead reagents against murine and human leukocyte antigens, or second-step reagents against various primary antibodies. These colloidal antibody-bound particles are said to be so small (50 nm) that

ADVERTISEMENTS

DNA SEQUENCING

Using an ABI 373 Automated Sequencing System
£60/sample

- PCR products
- M13 inserts
- plasmid inserts
- single stranded DNA
- double stranded DNA
- solid phase sequencing
- 300-400 bases of sequence

Please contact us for more information and a preparation protocol

Molecular Medicine Unit

King's College School of Medicine & Dentistry

Tel: 071-326 3126 Fax: 071-733 3877

Reader Service No.10

OLIGONUCLEOTIDES

Synthesised, purified and shipped in 3 days

£2.50 per base

(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit

King's College School of Medicine & Dentistry

Tel: 071-326 3126 Fax: 071-733 3877

Reader Service No.11

Unit 308, Technopole
Trafalgar Street
Newcastle upon Tyne
NE1 2LA
UK

Tel: (IDD 44) 091 201 2179
Fax: (IDD 44) 091 201 2191/230 3222

epitope
custom
peptides

- Peptide Synthesis**
\$12/residue 30-50mgs
- Peptide Conjugation**
KLH, BSA, Biotin
- Ab. Production & Purification**
PABs from \$950
- Protein Sequencing**
\$20/residue

Reader Service No.8

PACEMAKER®

Custom peptide, conjugate and antibody production services

A range of quality services based on experience and expertise.

Peptide purities at >70%, >80%, or >95%; quantities from 1mg.

Hplc and mass spectrometric analysis for all peptides.

Conjugates prepared using carefully optimised procedures.

Protocols to comply with individual requirements.

Antibodies of defined specificity for utilisation in a variety of biochemical and immunological techniques.

Prices from \$105 per peptide, \$75 perconjugate.

For a brochure, further information or advice please contact the experts at:

AFFINITY Research Products Ltd
GPT Business Park, Technology Drive, Nottingham,
NG9 2ND, United Kingdom.
Tel: (+44/0)-602-436100; Fax: (+44/0)-602-436300

Reader Service No. 9

they do not affect cell function or viability. The separated cells can be used for further studies, cell culture or flow cytometry. Separation purity is said to be greater than 95 per cent and yields are typically in excess of 90 per cent.

■ Plates

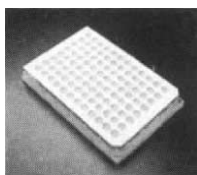
Xenopore's Xenoprobe **plates with streptavidin covalently attached to the surface** provide a useful format for carrying out hybridization assays and quantitative PCR (*Reader Service No. 116*). The covalent linkage is said to provide stability of the immobilized molecules while retaining the ability specifically to attach to biotinylated molecules. The plates and strips are available in two forms: clear for colorimetric assays and opaque black and white for fluorescent and chemiluminescent assays.

New from Nunc are FluoroNunc **modules** and MicroWell **plates** intended for use in FIA or IFMA fluorescence assays (*Reader Service No. 117*). The FluoroNunc resin used is said to minimize background interference while, at the same time, maximizing light reflection. Five formats are available: solid, 96-well C-bottom MicroWell plates in white or clear; modules in frames; eight C-bottom wells/strip in white; 12 C-bottom wells in clear; and 16 flat-bottom wells in black — all of which are available either in PolySorp or certified MaxiSorp surfaces.

To meet the needs of new analytical methods such as the use of luminescent markers and DNA hybridization assays, Labsystems has redesigned its range of **microplates and strips** (*Reader Service No. 118*). The 96-Well Cliniplates are now being manufactured with a new white pigment and binding surface. The opaque pigment is said to prevent light transmission from well to well so samples with a very high light output will not affect adjacent wells. Reflection inside the well is also said to be moderated — improving the signal-to-noise ratio. The

8- and 12-well microstrips are being offered with a choice of three binding characteristics: untreated Universal Binding white and black microstrips for use where no binding is required or for use with hydrophobic and other molecules that bind to untreated polystyrene; Enhanced Binding white and black microstrips which have been treated to improve the binding of more hydrophilic molecules; and streptavidin-coated microstrips which enable the binding of biotinylated molecules. White and black 96-well Cliniplates are also sold with the Universal and Enhanced Binding features: Cliniplates coated with streptavidin are available on request.

Millipore's new line of **opaque, 96-well filtration plates**, which are compatible with direct microplate readers, has been developed with the MultiScreen bioassay system in mind (*Reader Service No. 119*). This assay system enables an entire assay to be carried out within a single microplate. The new opaque plates are designed to prevent

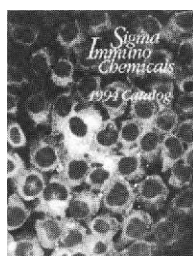


Opaque 96-well microplates.

interference between individual sample wells. By eliminating the need to transfer samples to scintillation vials (results can be read from the individual wells directly), this approach is said to reduce the amount of scintillation fluid used and, hence, waste, and to enhance the reproducibility and accuracy of assays by eliminating some of the typical causes of error associated with more conventional assays. The new plates are available with different types of Millipore filters, including glass-fibre filters for cell harvesting applications, Durapore membranes for assays requiring low-protein binding media, and nitrocellulose membranes for studies on immobilized biomolecules.

■ Catalogues/guides

Sigma's 1994 **immunochemicals catalogue** lists more than 1,600 antibodies, growth factors, enzyme substrate and buffer tablets and related products — over 170 of which are new listings (*Reader Service No. 120*). New products for analytical cytology include a line of mouse CD marker antibodies labelled with fluorescein isothiocyanate, phycoerythrin or Quantum Red. Among the new items for cell biology are antibodies to β -actin, dystrophin, golgi proteins, pan cadherin, plectin and syntaxin. Twenty new growth factors and antibodies to growth factors are listed. Also offered are pre-weighed HEPES, MES and MOPS buffer pouches.



On sale at Sigma.

Maxim Biotech's 1994 **catalogue** features products for molecular biology, immunohistochemistry and biochemistry (*Reader Service No. 121*). Products for molecular biology include primers for DNA amplification and probes for hybridizations (the products cover human inherited diseases, HLA polymorphism, cytokines, oncogenes, drug-resistant genes and viral, bacterial and parasitic diseases), *in situ* hybridization/detection kits and custom oligonucleotide synthesis services. Antibodies and positive control slides, a custom antibody specificity test, immunohistochemical kits, western blot kits and custom peptide synthesis services should interest researchers using immunohistochemical methods. Assorted substrates and buffers and nonradioactive oligonucleotide coupling reagents and columns complete the line-up.

Two new texts are available from ICN Biomedicals: a **reference guide to new growth factors and antibodies** and a **catalogue of immunochemicals** (*Reader Service No. 122*). The new 84-page guide



ICN reference guide.

contains technical information on 161 growth factor and growth factor antibodies, including information about source, form, activity, mode of action, applications and technical comments for each growth factor. In addition, the guide lists over 250 references. The new 224-page catalogue lists monoclonal and polyclonal antibodies, antibody conjugates, cytoskeletal proteins, immunohistology products, blood proteins and ancillary reagents.

Endogen's new 1994 product catalogue features over 150 **reagents and ELISA kits for immunological research** (*Reader Service No. 123*). New products include ELISA kits to quantify human IL-2, human IL-4, human IL-10, mouse soluble ICAM-1 (CD54) and mouse interferon gamma; a new line of recombinant mouse cytokines; a variety of monoclonal antibodies to cell-surface markers such as anti-mouse LFA-3 (CD58), anti-mouse IL-2R beta and anti-human PECAM-1 (CD31). The new Ultra-Sensitive ELISA kit line includes ELISAs that are said to measure human IL-2 and IL-6 at the femtogram level of sensitivity.

NBS Biologicals announces its new range of **cytokines and growth factors** (*Reader Service No. 124*). The new line-up includes recombinant human IL-11 and IL-13, HIV/HCV antigens, immunohistochemical reagents and associated substrates. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more details fill in the reader service card bound inside the journal.

ADVERTISEMENT

NEUROSCIENCE ANTIBODIES

- ✓ Choline Acetyltransferase
- ✓ Glutamate Receptors
- ✓ Neurofibrillary Tangles
- ✓ NMDAR1-New
- ✓ Oligodendrocytes
- ✓ Tau
- ✓ Tyrosine Hydroxylase

For a complete listing of our products please call, fax or write us today.

800-437-7500



CHEMICON International, Inc.

28835 Single Oak Drive, Temecula, CA 92590 USA
Phone 909-676-8080 • Fax 909-676-9209

Reader Service No.6